CLINICAL INVESTIGATION

Considerations in the design of clinical trials for relapsing multiple sclerosis

Clin. Invest. (2012) 2(11), 1073–1083

Trials in relapsing multiple sclerosis have been complicated by difficulties in defining the two principal significant clinical events (relapses and disability progression) that contribute to the devastating permanent neurological impairment characteristic of end-stage multiple sclerosis. There are further difficulties measuring the impact of these events, and the commonly used MRI end points used in Phase II studies have a complex relationship with both events and permanent neurological disability. Current designs favor the parallel-group, randomized, placebo-controlled study with annualized relapse rates as the principal clinical outcome in Phase III studies. As therapies have been developed, problems have arisen with changing trial populations with lower annualized relapse rates that vary both between studies and during the course of studies, making comparing therapies across studies difficult. Although novel designs have had very limited use thus far, active comparator trials, adaptive designs and targeted therapies/biomarkers have the potential to streamline therapy development in an increasingly crowded therapeutic marketplace.

Keywords: annualized relapse rate • clinically isolated syndrome • MRI • progression • relapse • relapsing multiple sclerosis • surrogate markers

Richard Nicholas*[1] & Tim Friede[2]
[1]Imperial College Healthcare NHS Trust, Charing Cross Hospital, Fulham Palace Road, London, UK
[2]Deptartment of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany
*Author for correspondence:
Tel.: +44 20 8383 0675
Fax: +44 20 8383 0584
E-mail: r.nicholas@imperial.ac.uk

Multiple sclerosis (MS) is a chronic inflammatory disease of the CNS. Age of onset is broad, peaking at between 20 and 40 years [1]. In Europe and North America, prevalence is 1/800 people, with an annual incidence of 2–10/100,000. These combine to make MS the most common cause of neurological disability in young adults [2,3]. Apart from a minority of people with 'aggressive' MS, the median disease course can be up to 50 years [4]. MS is a complex condition for therapy development due to widely varying outcomes that occur in the absence of therapy, the length of the disease course and the difficulty relating shorter-term relapse outcomes with later progression [5]. These factors have resulted in limited evidence as yet of the true benefit of currently licensed therapies over the long term. Trials, thus far, have primarily targeted relapsing MS (RMS), principally becasuse treatment early in the disease may prevent fixed future disability, although trials are increasingly targeting progressive MS. Classically, placebo-controlled, parallel-group studies have been used but new trial designs are emerging in small-scale studies. Therapies are increasingly effective at reducing relapse frequency, but an increasing number of licensed therapies that impact the available population for trials and make differentiation of new therapies more difficult mean that drug development efficiency needs to be improved.

Defining RMS

Treating a RMS in the context of trials requires both a correct diagnosis of MS and a correctly diagnosed relapse. In terms of diagnosis, aside from brain biopsy

FUTURE SCIENCE part of fsg

targeting an affected area of the brain, which cannot be justified in the vast majority, no definitive diagnostic test has yet been identified and a criteria-based diagnosis is used that relies on clinical assessment with ancillary tests, principally MRI. These tests are only of use in the context of an appropriate clinical history, with false-positive rates of diagnosis of 5% [6,7]. In general, MS is viewed as a single disorder with clinical variants. However, it may consist of several related disorders with differing immunological, pathological and genetic features [8,9]. Furthermore, 'demyelinating syndromes' considered distinct from MS have emerged; most notable are Devic's disease (neuromyelitis optica), relapsing optic neuritis and relapsing myelitis, which are believed to be predominantly due to antibody-mediated immunological damage. A consensus is emerging that optimal therapies may differ in these conditions compared with MS [10].

Three clinical presentations dominate and are expressions of key clinical elements of the disease; relapses and progression. Two presentations, relapsing remitting MS (RRMS) and secondary progressive MS (SPMS), can have relapses present when the condition is deemed RMS. They are:

- RRMS: characterized by episodes of neurological dysfunction and relapses, interspersed with periods of stability. In 90% of people, early disease is RRMS. Although some have a benign course over many years, most develop secondary progressive disease, usually 6–10 years after onset and in a few relapses persist;
- Primary progressive MS: in 10% of patients there is a continuous deterioration in neurological function from the outset;
- SPMS: follows RRMS when relapses may or may not continue but a continuous deterioration in neurological function occurs.

As treatment has been developed for MS, more importance has been attributed to the first demyelinating event or 'clinically isolated syndrome' (CIS), a single episode of neurological dysfunction lasting for greater than 24 h, which is the beginning of MS in the context of further clinical or MRI activity [11]. The outcome in CIS without therapy, given that a significant number have no further clinical or MRI activity, is much improved over RMS [12] and therapies have targeted this condition aiming, as with RMS, to prevent the later onset of disability by targeting future relapses. This is the population that enter RMS clinical trials; as a result, a wide spectrum of outcomes would be expected.

Measuring clinical events in MS

Relapses

The operational definition of a relapse is defined as a period in which a person with MS experiences an 'acute worsening of function' that lasts for at least 24 h, usually lasting for several days or weeks, followed by an improvement that lasts for at least 1 month [13]. 'Acute worsening of function' implies the development of new symptoms or previous symptoms returning without obvious reason [101]. Sometimes there is full recovery following a relapse, but quite often a person is left with residual problems and there can be no recovery. People can experience a short-term worsening of symptoms lasting minutes, hours or even days, which is not due to a relapse. An increase in body temperature due to an infection or following exercise can do this. Thus, defining relapses is not an exact science. Despite this, a person's previous history of relapses decides their eligibility for disease-modifying treatments [102]. More recently in trials, relapses have relied on objective evidence of deterioration on scales [14], but otherwise require documentation at the time by a trial physician. The annualized relapse rate (ARR), is the total number of relapses experienced by the cohort divided by the person/years at risk, and is the most commonly used summary measure of relapses.

Progression

Defined as the progressive accumulation of disability, assessed prospectively or retrospectively, that is persistent and is not due either to relapse-related impairment or temporary changes resulting from variations in disease or from the insensitivity of the method of measurement. This definition is problematic for trials for a number of reasons [15].

Any deterioration has to be persistent and permanent. Trials use definitions based on shorter time frames to reduce the time over which trials are conducted, potentially enhancing the impact of interventions on scales that measure progression due to residual effects of relapses.

Relapse deterioration is measured using the same scales that are used in progression. Therefore, in RMS, the impact of a relapse must be excluded to truly measure progression. A step deterioration after a relapse is not progressive accumulation of disability and is consistent with RMS. Thus, two measures of disability in time are required to demonstrate progression in the absence of relapse.

Correlating clinical events: relapses & progression with outcome

Trials in RMS can demonstrate that therapies can impact MRI activity and relapses can show an impact

on short-term progression outcome [14]. The problem still remains of how relapses interact with subsequent progression. Figure 1 implies a direct relationship in that relapses drive progression. However, in the absence of therapy it may be that only relapses in the first 2 years have an impact on subsequent progression [5], and only MRI T2 lesion activity in the first 5 years correlates with the development of SPMS [12]. Therefore, the merits of treatment after the early phase of the disease is unclear, but this will depend on how long the 'window of treatment opportunity' exists for an individual [16]. However, if everyone is treated early on, one will end up treating a condition with a median disease length of up to 50 years [4], and the earlier treatment is initiated the more likely one is to treat those who will not develop significant disability. Treating the disease early will also produce greater impact from therapy-related side effects given that the patient, if treated early, will have few disease-related effects [17]. This has to be balanced against the finding that once relapses are not evident, a 'relapse-modifying' therapy is not effective [16,18–20].

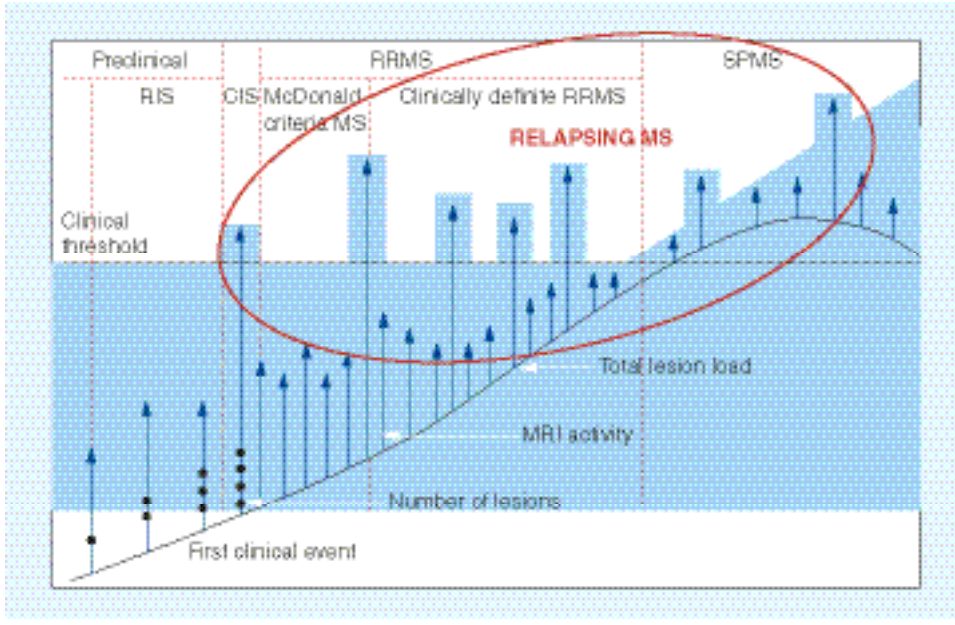

Figure 1. The clinical disease course in MS illustrating the time frame when people are considered to experience relapsing MS.
CIS: Clinically isolated syndrome; MS: Multiple sclerosis; RIS: Radiologically isolated syndrome; RRMS: Relapsing remitting multiple sclerosis; SPMS: Secondary progressive multiple sclerosis.
Adapted with permission from [17].

Thus, we have a condition that expresses itself clinically in two ways – relapses and progression – but it is not clear if they will occur and it can take years for these events to impact the individual. This lack of certainty remains a dilemma in MS therapeutics and means that there is an argument for not utilizing imperfect therapies with side effects in particular subjects who, with current knowledge, cannot be identified before disability accumulates [17].

Measuring disease activity

To measure disease activity, it is necessary to link the underlying pathology with surrogate outcomes, and hence to the clinical expression of the disease as measured using scales. This enables assessment of a therapy's effect on the disease (Figure 2). This link is vital to clarify if potential treatment is to be used early on, before irreversible disability develops.

■ Pathology

It is widely accepted that the development of CNS inflammatory lesions [21] is responsible for acute relapses in RMS and is associated with changes in the blood–brain barrier permeability, as determined by the presence of gadolinium (GAD)-enhancing lesions and subsequently with T2 lesions on MRI scans (Figure 2A). Axonal loss is believed to be the major process underlying the accumulation of irreversible disability during both relapsing and progressive phases of MS [22].

■ Surrogate markers for trials: MRI

MRI has revolutionized the diagnosis of MS and is the most commonly used biologically plausible outcome in randomized controlled trials; however, its utility as a surrogate marker is variable [23]. MRI T2 and GAD-enhancing lesions correlate with inflammatory pathology in early MS [23], but their correlation with subsequent disability is poor [24]. Newer MRI-based measures such as atrophy remain unproven in clinical trials [23,25].

■ Quantifying clinical events

Relapses using in-person examination to determine objective change are easier to measure compared with progression. However, when relapses are defined using confirmatory objective scales they become less frequent [14]. This is because quantifying change in MS remains a challenge due to the plethora of impacts the disease has on the person's physical, psychological and social function. It is very difficult to adequately

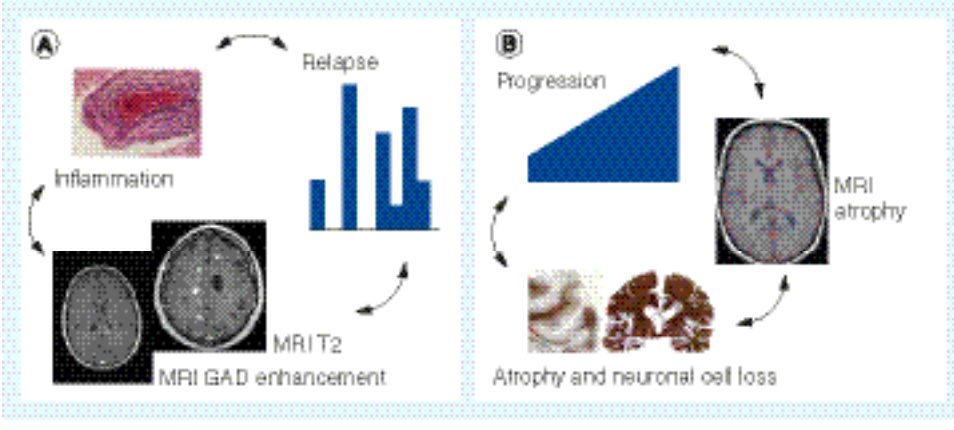

Figure 2. Proposed relationships between pathology, imaging and clinical course in relapsing and progressive multiple sclerosis. (A) Inflammation is pathologically inter-related with the clinical event relapse and T2- and GAD-enhancing lesions on MRI. **(B)** Brain atrophy and neuronal cell loss is pathologically inter-related with clinical progression, which is associated with atrophy on MRI; however, the strength of the relationships is stronger for relapses than it is for progression.
GAD: Gadolinium.
(B) Data taken from [22].

express a disability framework that is consistent and coherent when it can range from transient impaired colour vision to complete paralysis. In addition to causing neurological dysfunction, MS can also affect mood, fatigue, pain, work capability and a person's social position.

Measurement scales remain in their infancy in terms of understanding what change means [26], but the scales currently used widely in trials, such as the expanded disability status score (EDSS) and MS functional composite [27], have known limitations [28]. The EDSS ranges from 0 (no disability) to 10 (death from MS) in half-point increments. Lower scores (0–4) reflect specific neurological impairments and disability; higher scores reflect reducing levels of mobility (4–7) and upper-limb and bulbar function (7.0–9.5). The scale is nonlinear and patients spend different periods of time at defined levels. This impacts the population suitable for a study and has been criticized for indicating change poorly, for emphasizing neurological examination and mobility, and for failing to reflect other disabilities. Despite this, overall change in the EDSS does impact health utilities [29], but ultimately it is not sufficiently sensitive enough to measure progression over the time frames of RMS trials [30]. The MS functional composite was developed as a clinically meaningful outcome measure for trials, but again has similar limitations to the EDSS, is difficult to interpret, and although used, it is not currently a validated primary outcome in MS trials.

Current designs

Traditionally, trial structure to develop therapeutic compounds involves two phases; a learning phase and confirming phase. The approach for a particular disease further refines this approach determining the outcomes and time frames required at each stage to show a therapy is safe and effective. In MS, and particularly RMS, these parameters have been determined by regulators [31]. The current therapeutic approach to MS involves three arms: symptomatic treatment, treatment of relapses and treatment aimed at modifying the course of the disease. The strategy of design of current trials is described in Box 1.

■ The parallel-group, randomized, placebo-controlled study in RMS trials

In RMS, treatment has two aims: preventing or delaying the accumulation of disability and preventing or modifying relapses and the sequelae of acute relapses.

The range of relapse intensity and frequency, as well as recovery is highly variable, thus trials need to be long and large enough to overcome this inter- and intra-individual variability. It is acknowledged by regulators that the relationship of relapses to the accumulation of disability is unclear and that there are no surrogates for evaluating the progression of disability.

Short-term stabilization of the condition has been seen and can be attributed to the regression to the mean [32], the placebo effect or the natural history [33]. This effect means that regulators favor superiority trials either against placebo or an active comparator. Other designs considered are three-arm placebo and active-comparator, add-on designs, or placebo-controlled in those refusing standard therapies (Box 1). Regulators have also recommended including standard therapies that are not powered to be compared with the experimental drug, which, in turn, is powered against placebo. This enables an understanding of the impact of standard therapy in the treated cohort given the changing population entering trials over time.

■ End points in RMS trials

Primary efficacy end points preferred by regulators remain clinically measured prevention or delay in

the accumulation of disability [31]. Relapse rates may also be used as primary end points and are a common choice in recent trials [14,34], with relapses being predictive of disease progression as described in 'Correlating clinical events: relapses and progression with outcome' above. As secondary end points, disability and relapses should be measured, if not used as a primary end point. In addition, MRI is used using T2 sequences. Other measures of impairment can be utilized, such as cognition, fatigue and ambulation, since there is no single end point that captures the totality of the disease's impact on the individual [35]. Therefore, one might argue that the most promising approach lies in the integration of currently used and possibly some new end points [36]. In addition, different end points are suited to different stages of drug development and to the proposed action of the drug. The predictive value of various outcomes has recently been reviewed [36].

Despite its limitations, the EDSS is recommended as at least one of the measures to enable comparison with future measures. Mean change is not sufficient, as it is not a linear scale, but should be prespecified and any deterioration be confirmed by the same physician at least 6 months apart – note this is not long enough to confirm a true effect on progression. Relapses need to be defined blinded to therapy in terms of their occurrence, for example, time of beginning, time of ending, minimum duration to qualify as a relapse, maximum time elapsed between two symptoms to qualify as a single relapse, and severity. Effects on relapses should be shown for at least 2 years, though an effect can often be demonstrated at 1 year and the ARR is an acceptable parameter to assess relapses. Currently, MRI is not validated as an outcome for pivotal trials. However, it is suitable for use in exploratory studies – this includes GAD-enhancing and/or new or enlarging T2 lesions. Newer measures that have been related to tissue loss include T2 lesion load, chronic T1 weighted hypointensity (chronic black holes) and brain atrophy measures, but remain unproven in clinical trials.

Different subgroups of patients are suitable for certain trials defined usually by the EDSS score and relapse frequency; however, extrapolation from one subgroup to another is not possible.

Modelling relapses, progression & MRI outcomes

Relapses are commonly modelled as Poisson counts with the scaled event rate being the ARR. To account for between-patient heterogeneity, adjustments to the Poisson model and use of the negative binomial distribution have been suggested more recently [37]. However, all these models assume constant relapse rates, an assumption that was challenged by a recent meta-analysis [38]. Sustained progression in EDSS is usually analyzed as a time-to-event end point using nonparametric techniques as Kaplan–Meier curves and log-rank test or semiparametric methods such as Cox regression, which allows adjustment for potential confounders [39]. However, more sophisticated methods for the analysis of EDSS data have been suggested including a Markov transitional model [40], the use of random effects Markov models to predict progression [41], and their extension to processes evolving continuously over time with discrete observation time points [42]. MRI volume measurements, such as

Box 1. The strategy of design of current trials.

Pharmacodynamics

Animal models to understand mode of action
Dose finding and side effects in healthy volunteers and patients

Exploratory studies

Subjects with high chance of showing an effect
MRI as a primary outcome using a parallel double-blind design for ≥4 months
Evaluate disability, relapses and clinically meaningful outcomes
Evaluate safety and the absence of worsening MS

Confirmatory studies

Main efficacy end point depends on the goal of the treatment
Large scale and long enough to have a substantial proportion of patients suffering relapses or showing progression of disability
Usually double-blind parallel study 2 years in duration
Superiority trial versus placebo or any available single therapy
If not clear from earlier studies a number of doses should be tested
To understand the long-term course of patients under treatment, an extended open-label follow up should be performed

- Option 1: active-control parallel-group trials comparing new treatment to approved treatment, ideally three-arm studies with placebo, test product and active control
  Ensure no selection of patients having shown nonresponse (or suboptimal response) to previous therapy
- Option 2: compare the new treatment with placebo in a short-duration trial and thereafter to switch placebo patients to a predefined active treatment or randomize them to the experimental product or a predefined active treatment
- Option 3: if the product is to be used in combination, a placebo-controlled trial of 2–3 years duration as an add-on treatment

MS: Multiple sclerosis.
Data taken from [31].

T2-lesion burden, are sometimes transformed by taking the cube root prior to analysis [43], whereas MRI lesion counts are statistically modelled by the so-called negative binomial distribution that accounts for some heterogeneity in lesion incidence between patients [44–46].

In addition, the question on how to deal with excess numbers of zeros has very recently garnered some attention [47]. The authors considered a zero-inflated negative binomial model, which extends previous work by including an additional mixture component for the zeros. Weaknesses in the design, analysis and reporting of clinical trials in MS have very recently been highlighted by Signori *et al.* although an overall improvement over time was noted [48].

Emerging trial designs

Newer trial designs have started to emerge in RMS, but they have tended to be applied in small-scale circumstances and have, as yet, not been applied in large-scale trials or exclusively in the absence of traditional approaches. Active comparator designs have been used, but together with placebo trials, not alone. Fingolimod, the most recently licensed compound, has been tested against Avonex®, a standard therapy [49]. Early adaptive designs have started to appear in early Phase II trials in RMS [50] with an appreciation that our limited understanding of drug doses required means that the flexibility that adaptation gives can improve the drug-development process. The development of daclizumab, a therapy based on its impact on a biomarker, has led to full-scale traditional Phase III trials [51,52].

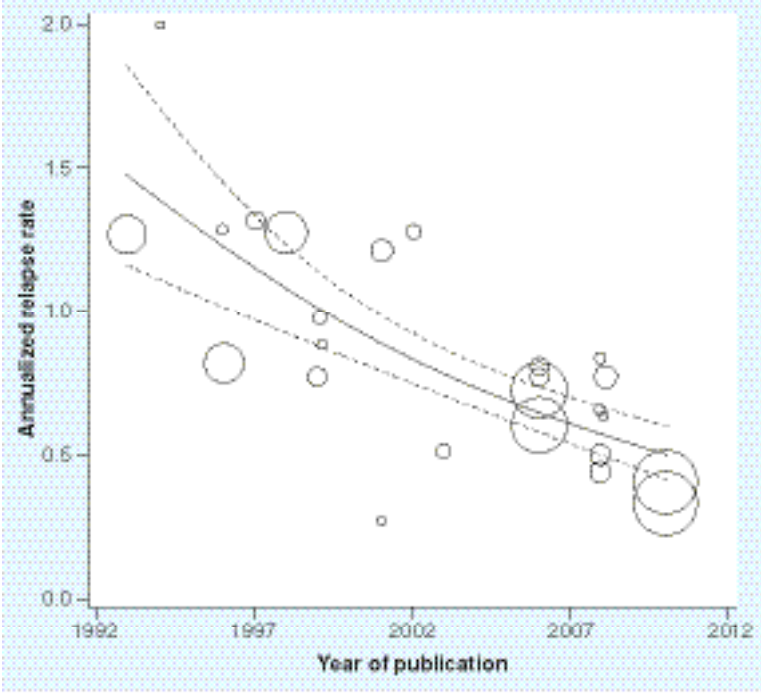

Figure 3. Annualized relapse rate observed in 26 trials against calendar year in which the papers were published. The size of the bubble indicates the size of the trial and is proportional to the sample size. The solid line shows the model values of the negative binomial regression and the dashed lines indicate 95% confidence intervals.
Reproduced from [42].

Issues with current designs

The current approach has started to pay dividends in terms of delivering therapies for RMS. Driven by large Phase III pharmaceutical trials, a world-wide infrastructure has emerged to deliver standardized protocols and definitions using short-term clinical and radiological outcome measures that have laid the foundation for an evidence-based approach to the treatment of MS [53]. An emerging understanding of the aims of trials in MS, and the licensing of therapies, together with the development of new statistical techniques offer the possibility of improving the efficiency of the current designs and of increasing their utility without impacting on safety.

Since the number of patients recruited into a trial impacts on cost, trial duration and number of participating centers, sample size calculation is a key component in the design of any clinical trial. Planning of a randomized controlled trial in RMS, with clinical relapses as the primary end point, sample-size calculations need to consider the assumed ARR in the control group, a factor accounting for between-patient heterogeneity, the effect size, the length of follow up, the desired power, and the significance level [54]. Some of these design aspects have been challenged over time by changes in MS trial recruitment.

Reducing placebo ARR with time

The relapse rate in subjects taking part in trials has dropped over the past two decades and in the context of developing therapies [54,55]. It seems likely that "this is not an indication that clinicians are seeing patients who are less sick or plagued by fewer relapses but is a result of changes in the trial environment" [55]. Most prominent is that earlier trials may have used more active patients. Current randomized, controlled trials recruit in the context of therapies developed by earlier studies and accepted by clinicians. This encourages the inclusion of milder patients in trials as more active subjects will be offered licensed therapies in preference to unproven approaches. This switch in selection clearly would lower the relapse rates. Other issues may arise from changes in the definition of MS as a result of the use of the McDonald criteria and the adoption

of precise definitions of relapse [14], eliminating false-positive relapses and potentially reducing relapse rates. The allowance of rescue therapy and subsequent exit from trials again could contribute to a lowering of in-trial relapse rates [54]. If the rescue therapy hypothesis really is the cause of these lowered ARRs, much more careful attention must be paid to drop-outs in interpreting contemporary clinical trials.

Comparing therapeutic effects on ARR

Variation in placebo relapse rates makes the interpretation of relative effects from various trials difficult. A recent review of the earlier therapies found an absolute risk reduction from 0.15 to 0.43 relapses per year corresponding to a relative risk reduction of 18–34% [53]. Compared with recent placebo-controlled trials of natalizumab [56], fingolimod [14] and cladribine [34], which reported relative risk reductions in the range of 55–68%. However, when expressed as absolute risk reduction, the treatment effect sizes (0.18–0.5 fewer relapses per year) are in fact quite similar to those seen with the first-generation therapies [53]. It is the absolute reduction in relapses that impacts the number needed to treat; thus, though more effective in terms of relative risk reduction, the cost per relapse saved in fact increases. However, the variation in the placebo arms makes using number needed to treat problematic. These issues are important in cost–effectiveness analyses of treatments, and have contributed to the problems with fingolimod accessing the British market and gaining NICE approval [103].

Variation in trial ARRs

In the face of reducing trial ARR, there remains considerable variation about this trend (Figure 3), which results in uncertainty in the planning of a new trial. Thus, for recent studies, estimates of AAR have been 0.7 whereas the observed ARR was 0.4 [14]. This makes it difficult to calculate an appropriate sample size. If too large, the cost is increased unnecessarily or if too small, potentially a trial could fail when the therapy is appropriate.

Reducing ARR within studies

While the ARR is commonly assumed to be constant over follow-up times, a systematic review very recently challenged this assumption with placebo ARR dropping with study follow-up time (Figure 4) [54]. This is due, in part, to the regression to the mean phenomenon, which was described in MS [32] and is well known in other conditions such as epilepsy [57]. Reducing within-study placebo ARR affects the planning of trials, in that data from short-term trials might not be appropriate for planning of longer studies. Furthermore, whether the treatment effects are constant over time or vary with time is of increasing interest in clinical practice with the recent review suggesting they are not constant [54]. Increasing options for therapy, short trial designs and a drive to identify effective therapies as early as possible make it important to recognize whether a treatment effect is waning or increasing over the course of a study. It is further complicated by the fact that the relapse rate reduces by 17% every 5 years as the disease course progresses [33].

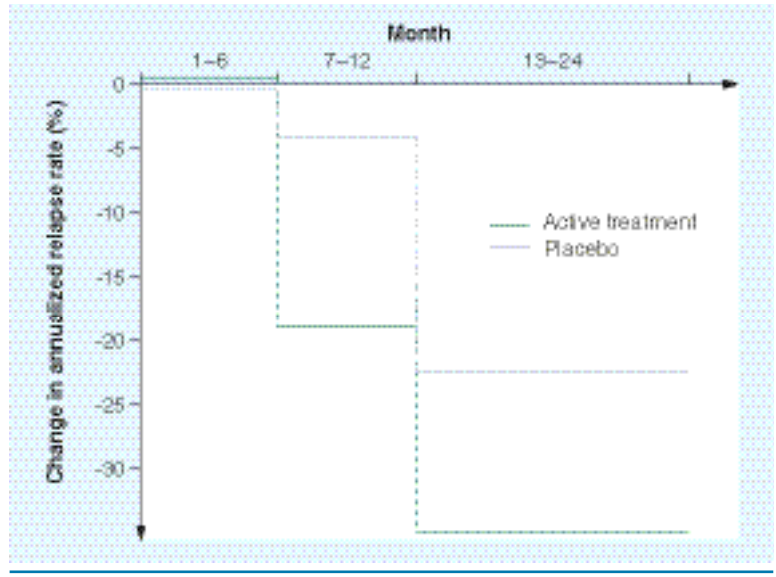

Figure 4. Decrease in annualized relapse rate (in percentage) between the early and late time periods and how that contributes to a differential effect of active treatment and placebo by the end of the trial.
Reproduced from [47].

Subgroup analysis to identify comparable populations

Comparison of the affects of therapies on ARR remains a problem when apparently highly active groups are sub-selected from these newer trials to enable comparison with historical studies [48]. These analyses feed directly into reimbursement [103], despite the credibility of claims from subgroup analyses usually being low [58]. A concern remains that these purportedly highly active groups chosen from a more benign trial cohort are not the same as those with more aggressive disease who access standard therapies in preference to trials.

Placebo-controlled trials in RMS

Trials using a placebo are becoming increasingly untenable when therapies are available. In environments where treatment is available, individuals should be offered therapy and a robust consenting process

is required [59]. This changes the trial population as those taking part will have refused therapy or will have failed standard therapies. Though comparator studies are gaining favor, there is little work on add-on studies as combination therapy has been associated with increased side effects in the past.

Heterogeneity in MS

The variability of MS as a disease is exacerbated by the potential for inclusion of different demyelinating syndromes that are difficult to recognize especially as treatment is instituted earlier. Higher heterogeneity, will increase the number required to show an effect. Early stage trials in conditions such as CIS will lead to the increasing incorporation of subjects with benign disease, again increasing the sample size (Figure 1).

Surrogate end points

There is an increasing need for early outcomes to predict treatment effects of candidate therapies and thereby to facilitate screening. Surrogate end points have, by definition, two distinct properties [60]. First, they correlate on an individual level with a clinical outcome. Second, treatment effects as measured on the scale of the surrogate outcome are correlated with treatment effects on a clinically relevant outcome. Generally, individual level correlation is poor for early outcomes with longer term clinical outcomes with MRI outcome offering little in addition to clinical outcomes [61]. Treatment effects on MRI outcomes are associated with treatment effects in terms of relapses [62–64], and MRI sequences have been claimed to give additional insights into the mechanisms of drug action [65].

Future perspective

Redefining MS

Current trial designs are tied to the clinical features of MS and their limitations. With our emerging understanding of MS, in future we may be able to reclassify MS and thus redesign our trial approach – at the moment this is theoretical but developing techniques could enable this to be a reality in the future.

Active controlled trials

As therapies in RMS become available, the ethics of performing placebo-controlled clinical trials become more difficult and, as has occurred in many therapeutic areas, it is no longer practical to compare with placebo. The complexity of RMS and the lack of clear evidence of an effect on a progressive outcome make an active controlled trial's superiority and noninferiority more difficult as they are usually performed when there is a known therapeutic effect of the active treatment against placebo. They have been used, but in association with placebo trials and not yet in isolation.

Adaptive designs

Adaptive designs are seen as one way of increasing the efficiency of clinical research programs by combining previously separate learning and confirming phases into a single trial [104]. In Phase II and III clinical trials, sample size re-estimation, dose selection, and subgroup selection are the adaptations of greatest interest. Their utility for application in MS has been investigated [24,54,66] and some of these adaptations have been applied in RMS [50] in a small way, but the benefits of such an approach are yet to be utilized in large-scale trials.

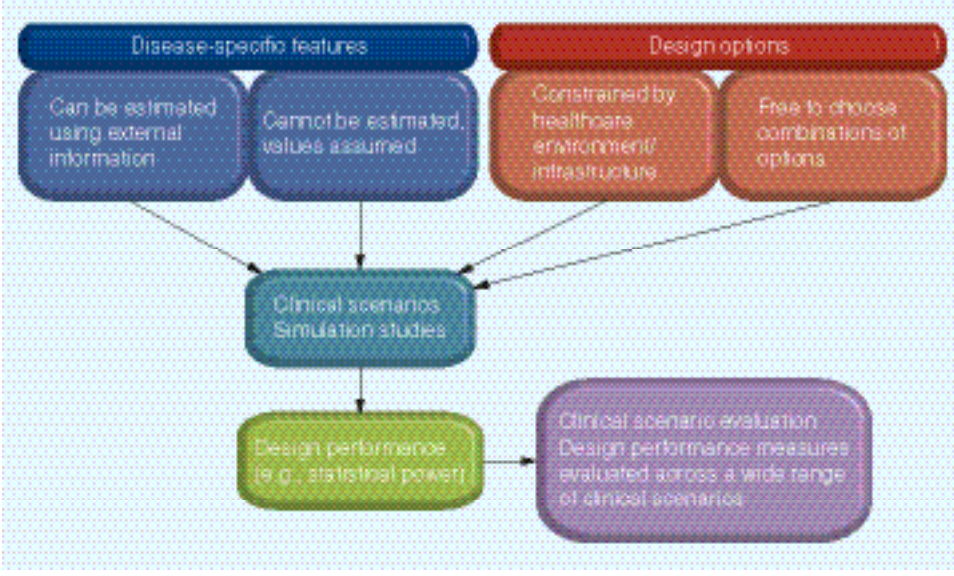

Figure 5. Clinical scenario evaluation framework.
Adapted from [62]. Reproduced with permission © DIA (2009).

Assessment of competing development strategies

The feasibility and utility of the above mentioned novel approaches to clinical trials in RMS can be assessed in comparison with more traditional clinical development programs through so-called clinical scenario evaluation [67,68]. This has been refined [69] and applied [24] in the context of MS trials. In a clinical scenario evaluation, competing strategies of clinical trial design or analysis are compared

Executive summary

Background
- Multiple sclerosis (MS) is a chronic autoimmune disease of the CNS. The outlook is highly variable in individual cases but disease progression overall makes MS the most common cause of disability in young adults.

Defining relapsing MS
- MS diagnosis is a criteria-based diagnosis. As well as three clinical presentations, of which two can be classed as being relapsing MS (RMS), there may be distinct immunological and genetic types, as well as potentially separate but related demyelinating syndromes. Clinically isolated syndrome, a precursor to RMS, further complicates the spectrum of subjects entering RMS trials.

Measuring clinical events in MS
- The principal clinical events that occur in MS are relapses and progression; they are related but the extent of their relationship and their impact on final outcome – permanent neurological impairment – is unclear.

Measuring disease activity
- Disease activity is usually measured in trials using scales to quantify disability/progression and frequency to measure relapses (e.g., annualized relapse rate [ARR]). Otherwise, disease activity is measured with the use of MRI scanning techniques, a biologically plausible outcome that is the principal surrogate outcome.

Current designs
- Designs are principally based on the parallel-group, randomized, placebo-controlled study using MRI in Phase II and clinical outcomes, practically ARRs, in Phase III trials.

Issues with current designs
- Current designs need revision as the known heterogeneity of MS itself is exacerbated in trials as the availability and use of partially effective treatments increases.
- The availability of treatments make it more difficult to justify the use of placebo in trials, which, in turn, impacts on trial populations. This has reduced the placebo ARR with time and during studies, increased trial-population heterogeneity. These changes make comparison of treatments across trials very difficult.

Future perspective
- Actively controlled trials and adaptive designs are beginning to be used in RMS trials. Advances in biomarkers and the targeting of therapies matched with a coherent assessment of competing approaches should increase trial efficiency, ultimately allowing RMS studies to test the effects of therapy on the more clinically and economically relevant progression outcomes.

through computer simulations for a set of clinical situations in terms of metrics, such as statistical power and sample size (Figure 5).

Biomarkers & targeted therapies

Across diseases, there is a move towards personalized medicine. In MS therapeutics, the concept or targeting treatments is emerging with the early success of daclizumab in Phase II trials [51] where a marker was utilized to identify responders [52]. Biomarkers potentially offer many benefits, helping to increase the efficiency of trials. Unfortunately beyond MRI, little has usefully emerged, though cerebrospinal fluid neurofilament analysis is currently being more fully evaluated [70].

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties. No writing assistance was utilized in the production of this manuscript.

References

1 Weinshenker BG, Bass B, Rice GPA *et al.* The natural history of multiple sclerosis: a geographically based study. 1. Clinical course and disability. *Brain* 112, 133–146 (1989).

2 Ford HL, Gerry E, Johnson M *et al.* A prospective study of the incidence, prevalence and mortality of multiple sclerosis in Leeds. *J. Neurol.* 249, 260–265 (2002).

3 Sloka JS, Pryse-Phillips WE, Stefanelli M. Incidence and prevalence of multiple sclerosis in Newfoundland and Labrador. *Can. J. Neurol. Sci.* 32, 37–42 (2005).

4 Degenhardt A, Ramagopalan SV, Scalfari A *et al.* Clinical prognostic factors in multiple sclerosis: a natural history review. *Nat. Rev. Neurol.* 5, 672–682 (2009).

5 Scalfari A, Neuhaus A, Degenhardt A *et al.* The natural history of multiple sclerosis: a geographically based study 10: relapses and long-term disability. *Brain* 133, 1914–1929 (2010).

6 Compston A, Coles A. Multiple sclerosis. *Lancet* 372, 1502–1517 (2008).

7 Miller DH, Weinshenker BG, Filippi M *et al.* Differential diagnosis of suspected multiple sclerosis: a consensus approach. *Mult. Scler.* 14, 1157–1174 (2008).

8 Compston A. Genetic epidemiology of multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* 62, 553–561 (1997).

9 Lucchinetti CF, Bruck W, Rodriguez M *et al.* Distinct patterns of multiple sclerosis pathology indicates heterogeneity in pathogenesis. *Brain Pathol.* 6, 259–274 (1996).

10 Sellner J, Boggild M, Clanet M *et al.* EFNS guidelines on diagnosis and management of neuromyelitis optica. *Eur. J. Neurol.* 17, 1019–1032 (2010).

11 Miller D, Barkhof F, Montalban X *et al.* Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis.

Lancet Neurol. 4, 281–288 (2005).

12 Fisniku LK, Brex PA, Altmann DR *et al.* Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain* 131, 808–817 (2008).

13 Poser CM, Paty DW, Scheinberg LC *et al.* New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann. Neurol.* 13, 227–231 (1983).

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

15 Ebers GC, Heigenhauser L, Daumer M, Lederer C, Noseworthy JH. Disability as an outcome in MS clinical trials. *Neurology* 71, 624–631 (2008).

16 Freedman MS. Multiple sclerosis therapeutic strategies: use second-line agents as first-line agents when time is of the essence. *Neurol. Clin. Pract.* 1, 66–68 (2011).

17 Nicholas R, Giannetti P, Alanousi A, Friede T, Muraro PA. The development of oral immune-modifying agents in the management of multiple sclerosis. *Drug Des. Devel. Ther.* 5, 255–274 (2011).

18 Coles AJ, Wing MG, Molyneux P *et al.* Monoclonal antibody treatment exposes three mechanisms underlying the clinical course of multiple sclerosis. *Ann. Neurol.* 46, 296–304 (1999).

19 Burt RK, Cohen BA, Russell E *et al.* Hematopoietic stem cell transplantation for progressive multiple sclerosis: failure of a total body irradiation-based conditioning regimen to prevent disease progression in patients with high disability scores. *Blood* 102, 2373–2378 (2003).

20 Naismith RT. Multiple sclerosis therapeutic strategies: start safe and effective, reassess early, and escalate if necessary. *Neurol. Clin. Pract.* 1, 69–72 (2011).

21 Lassmann H, Wekerle H. The pathology of Multiple Sclerosis. In: *McAlpine's Multiple Sclerosis (Volume 4).* Churchill Livingstone, London, UK, 557–599 (2005).

22 Reynolds R, Nicholas R, Roncaroli F, Radotra B, Gveric D, Howell O. The neuropathological basis of clinical progression in multiple sclerosis. *ACTA Neuropathologica* 17, 1211–1217 (2011).

23 Barkhof F, Filippi M. MRI–the perfect surrogate marker for multiple sclerosis? *Nat. Rev. Neurol.* 5, 182–183 (2009).

24 Chataway J, Nicholas R, Todd S *et al.* A novel adaptive design strategy increases the efficiency of clinical trials in secondary progressive multiple sclerosis. *Mult. Scler.* 17, 81–88 (2011).

25 Kapoor R, Furby J, Hayton T *et al.* Lamotrigine for neuroprotection in secondary progressive multiple sclerosis: a randomized, double-blind, placebo-controlled, parallel-group trial. *Lancet Neurol.* 9, 681–688 (2010).

26 Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol. Assess.* 13(12), 1–177 (2009).

27 Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. *Mult. Scler.* 5, 244–250 (1999).

28 Hobart J, Kalkers N, Barkhof F, Uitdehaag B, Polman C, Thompson A. Outcome measures for multiple sclerosis clinical trials: relative measurement precision of the Expanded Disability Status Scale and Multiple Sclerosis Functional Composite. *Mult. Scler.* 10, 41–46 (2004).

29 Naci H, Fleurence R, Birt J, Duhig A. The impact of increasing neurological disability of multiple sclerosis on health utilities: a systematic review of the literature. *J. Med. Econ.* 13, 78–89 (2010).

30 Daumer M, Neuhaus A, Herbert J, Ebers G. Prognosis of the individual course of disease: the elements of time, heterogeneity and precision. *J. Neurol. Sci.* 287(Suppl. 1), S50–S55 (2009).

31 Committee for Medicinal Products for Human Use (CHMP). Guideline on clinical investigation of medicinal products for the treatment of multiple sclerosis. Ref. EMEA/CHMP/EWP/561/98 Rev 1, London, UK (2005).

32 Martínez-Yélamos S, Martínez-Yélamos A, Ozaeta MG *et al.* Regression to the mean in multiple sclerosis. *Mult. Scler.* 12, 826–829 (2006).

33 Tremlett H, Zhao Y, Joseph J *et al.* Relapses in multiple sclerosis are age- and time-dependent. *J. Neurol. Neurosurg. Psychiatry* 79, 1368–1375 (2008).

34 Giovannoni G, Comi G, Cook S *et al.* A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. *N. Engl. J. Med.* 362, 416–426 (2010).

35 D'Souza M, Kappos L, Czaplinski A. Reconsidering clinical outcomes in multiple Sclerosis: relapses, impairment, disability and beyond. *J. Neurol. Sci.* 274, 76–79 (2008).

36 Goldman MD, Motl RW, Rudick RA. Possible clinical outcome measures for clinical trials in patients with multiple sclerosis. *Ther. Adv. Neurol. Disord.* 3, 229–239 (2010).

37 Wang YC, Meyerson L, Tang YQ, Qian N. Statistical methods for the analysis of relapse data in MS clinical trials. *J. Neurol. Sci.* 285, 206–211 (2009).

38 Nicholas R, Straube S, Schmidli H, Pfeiffer S, Friede T. Time-patterns of annualized relapse rates in randomized placebo-controlled clinical trials in relapsing multiple sclerosis: a systematic review and meta-analysis. *Mult. Scler.* 18, 1287–1293 (2012).

39 Phillips JT, Giovannoni G, Lublin FD *et al.* Sustained improvement in Expanded Disability Status Scale as a new efficacy measure of neurological change in multiple sclerosis: treatment effects with natalizumab in patients with relapsing multiple sclerosis. *Mult. Scler.* 17, 970–979 (2011).

40 Mandel M, Gauthier SA, Gutmann CRG, Weiner HL, Betensky RA. Estimating time to event from longitudinal categorical data: an analysis of multiple sclerosis progression. *J. Am. Stat. Assoc.* 102, 1254–1266 (2007).

41 Mandel M, Betensky RA. Estimating time-to-event from longitudinal ordinal data using random-effects Markov models: application to multiple sclerosis progression. *Biostatistics* 9, 750–764 (2008).

42 Mandel M. Estimating disease progression using panel data. *Biostatistics* 11, 304–316 (2010).

43 Li DK, Held U, Petkau J *et al.* MRI T2 lesion burden in multiple sclerosis: a plateauing relationship with clinical disability. *Neurology* 66, 1384–1389 (2006).

44 Sormani MP, Bruzzi P, Beckmann K *et al.* The distribution of magnetic resonance imaging response to interferon beta-1b in multiple sclerosis. *J. Neurol.* 252, 1455–1458 (2005).

45 van den Elskamp IJ, Knol DL, Uitdehaag BMJ, Barkhof F. The distribution of new enhancing lesion counts in multiple sclerosis: further explorations. *Mult. Scler.* 15, 42–49 (2009).

46 Aban IB, Cutter GR, Mavinga N. Inferences and power analysis concerning two negative binomial distributions with an application to MRI lesion counts data. *Comput. Stat. Data. Anal.* 53, 820–833 (2009).

47 Francois M, Peter C, Gordon F. Dealing with excess of zeros in the statistical analysis of

magnetic resonance imaging lesion count in multiple sclerosis. *Pharmaceut. Statist.* 11(5), 417–424 (2012).

48 Signori A, Baccino A, Sormani MP. The quality of reports of randomized trials in multiple sclerosis: a review. *Mult. Scler.* 18, 776–781 (2012).

49 Cohen JA, Barkhof F, Comi G *et al.* Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N. Engl. J. Med.* 362, 402–415 (2010).

50 Selmaj K, Li D, Stüve O *et al.* BAF312, a selective sphingosine 1-phosphate receptor modulator, effectively suppresses MRI lesion activity in relapsing-remitting multiple sclerosis: findings of an adaptive dose-ranging Phase 2 study. *Mult. Scler.* 17, S510 (2011).

51 Giovannoni G, Gold R, Selmaj K *et al.* A randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of daclizumab HYP monotherapy in relapsing-remitting multiple sclerosis: primary results of the SELECT trial. *Mult. Scler.* 17, S508 (2011).

52 Bielekova B, Howard T, Packer AN *et al.* Effect of anti-CD25 antibody daclizumab in the inhibition of inflammation and stabilization of disease progression in multiple sclerosis. *Arch. Neurol.* 66, 483–489 (2009).

53 Montalban X. Review of methodological issues of clinical trials in multiple sclerosis. *J. Neurol Sci.* 311(Suppl. 1), S35–S42 (2011).

54 Nicholas R, Straube S, Schmidli H, Schneider S, Friede T. Trends in annualized relapse rates in relapsing remitting multiple sclerosis and consequences for clinical trial design. *Mult. Scler.* 17, 1211–1217 (2012).

55 Inusah S, Sormani MP, Cofield SS *et al.* Assessing changes in relapse rates in multiple sclerosis. *Mult. Scler.* 16, 1414–1421 (2010).

56 Polman CH, O'Connor PW, Havrdova E *et al.* A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N. Engl. J. Med.* 354, 899–910 (2006).

57 Burneo JG, Montori VM, Faught E. Magnitude of the placebo effect in randomized trials of antiepileptic agents. *Epilepsy Behav.* 3, 532–534 (2002).

58 Sun X, Briel M, Busse JW *et al.* Credibility of claims of subgroup effects in randomized controlled trials: systematic review. *BMJ* 344, e1553 (2012).

59 Polman CH, Reingold SC, Barkhof F *et al.* Ethics of placebo-controlled clinical trials in multiple sclerosis: a reassessment. *Neurology* 70, 1134–1140 (2008).

60 Burzykowski T, Molenberghs G, Buyse ME. *The Evaluation of Surrogate End Points.* Springer, NY, USA (2005).

61 Daumer M, Neuhaus A, Morrissey S, Hintzen R, Ebers GC. MRI as an outcome in multiple sclerosis clinical trials. *Neurology* 72, 705–711 (2009).

62 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.* 65, 268–275 (2009).

63 Sormani MP, Stubinski B, Cornelisse P, Rocak S, Li D, De Stefano N. Magnetic resonance active lesions as individual-level surrogate for relapses in multiple sclerosis. *Mult. Scler.* 17, 541–549 (2011).

64 Sormani MP, Bonzano L, Roccatagliata L, De Stefano N. Magnetic resonance imaging as surrogate for clinical end points in multiple sclerosis: data on novel oral drugs. *Mult. Scler.* 17, 630–633 (2011).

65 Leist TP, Marks S. Magnetic resonance imaging and treatment effects of multiple sclerosis therapeutics. *Neurology* 74(Suppl. 1), S54–S61 (2010).

66 Riddell CA, Zhao Y, Petkau AJ. An adaptive clinical trial design for a sensitive subgroup examined in the multiple sclerosis context. *Mult. Scler.* 17, S35 (2011).

67 Benda N, Branson M, Maurer W, Friede T. Clinical Scenario Evaluation: a framework for the evaluation of competing development strategies. *Drug Development* 4, 84–88 (2009).

68 Benda N, Branson M, Maurer W, Friede T. Aspects of modernizing drug development using scenario planning and evaluation. *Drug. Inf. J.* 44, 299–315 (2010).

69 Friede T, Nicholas R, Stallard N *et al.* Refinement of the clinical scenario evaluation framework for assessment of competing development strategies with an application to multiple sclerosis. *Drug. Inf. J.* 44, 713–718 (2010).

70 Giovannoni G. Cerebrospinal fluid neurofilament: the biomarker that will resuscitate the 'Spinal Tap'. *Mult. Scler.* 16, 285–286 (2010).

Websites

101 Multiple Sclerosis Decisions. www.msdecisions.org.uk (Accessed 16 June 2012)

102 Association of British Neurologists. Revised Association of British Neurologists' guidelines for prescribing in multiple sclerosis (2009). www.theabn.org/abn/userfiles/file/ABN_MS_Guidelines_2009_Final%281%29.pdf (Accessed 16 June 2012)

103 NICE. Multiple sclerosis (relapsing-remitting) – fingolimod: final appraisal determination. http://guidance.nice.org.uk/TA/Wave20/71/FAD (Accessed 16 June 2012)

104 US FDA. Critical Path Opportunities List. www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/UCM077258.pdf (Accessed 16 June 2012)