

# Efficacy of Disease-Modifying Therapies in Relapsing Remitting Multiple Sclerosis: A Systematic Comparison

Mark S. Freedman<sup>a</sup> Bruce Hughes<sup>b</sup> Daniel D. Mikol<sup>c</sup> Randy Bennett<sup>d</sup>  
Brian Cuffel<sup>e</sup> Vamil Divan<sup>e</sup> Nicole LaVallee<sup>f</sup> Ahmad AL-Sabbagh<sup>d</sup>

<sup>a</sup>Department of Medicine (Neurology), University of Ottawa, Ottawa, Ont., Canada; <sup>b</sup>Ruan Neurology Clinic and Research Center, Des Moines, Iowa, <sup>c</sup>Department of Neurology, University of Michigan, Ann Arbor, Mich., <sup>d</sup>EMD Serono, Inc., Rockland, Mass., <sup>e</sup>Pfizer Inc., New York, N.Y., and <sup>f</sup>PROMETRIKA LLC, Cambridge, Mass., USA

## Key Words

Disease-modifying therapy · Evidence-based medicine · Interferon beta · Glatiramer · Natalizumab · Therapeutic gain

## Abstract

The treatment of relapsing-remitting multiple sclerosis (RRMS) has become more effective over the last decade with the advent of the currently available disease-modifying therapies (DMTs). Pivotal clinical studies differ in many characteristics, such that cross-comparisons of relative risk reductions are of limited value and can be misleading. Our objective was to compare the clinical efficacy of currently approved first-line DMTs in patients with RRMS, applying an evidence-based medicine approach. We reviewed all phase III pivotal trials of DMTs. Six clinical trials of Avonex<sup>®</sup>, Betaseron<sup>®</sup>, Copaxone<sup>®</sup>, Rebif<sup>®</sup> and Tysabri<sup>®</sup> in patients with RRMS were identified for analysis. Only randomized, placebo-controlled, double-blind studies were included. The clinical efficacy endpoints compared were: proportion of relapse-free patients at 1 and 2 years; annualized relapse rate at 2 years; proportion of progression-free patients at 2 years, and proportion of patients free of gadolinium-enhancing lesions at 1 year or 9 months. Based on these analyses, Betaseron, Rebif, and Tysabri show comparable effects, whereas for sev-

eral endpoints Avonex or Copaxone did not significantly differ from placebo. In the absence of head-to-head studies for all products used to treat RRMS, it still may be possible to compare treatment effects by applying evidence-based medicine principles.

Copyright © 2008 S. Karger AG, Basel

## Introduction

The treatment of relapsing-remitting multiple sclerosis (RRMS) has changed substantially with the advent of effective disease-modifying therapies (DMTs) that can alter the natural history of disease. DMTs have demonstrated the ability to reduce the incidence of relapses and the MRI activity thought to be linked to the underlying perceived inflammatory phase of the disease. They also retard disease progression, as assessed clinically by changes in a neurological rating scale called the Expanded Disability Status Scale (EDSS). There are currently 6 Food and Drug Administration (FDA)-approved treatments for RRMS in the United States. In 1993, the first DMT, interferon beta (IFN $\beta$ )-1b (Betaseron<sup>®</sup>), was approved for RRMS to reduce the rate and severity of relapses. This was followed by approval of IFN $\beta$ -1a (Avonex<sup>®</sup>) in 1996 and glatiramer acetate (Copaxone<sup>®</sup>) in 1997.

## KARGER

Fax +41 61 306 12 34  
E-Mail [karger@karger.ch](mailto:karger@karger.ch)  
[www.karger.com](http://www.karger.com)

© 2008 S. Karger AG, Basel  
0014-3022/08/0601-0001\$24.50/0

Accessible online at:  
[www.karger.com/ene](http://www.karger.com/ene)

Dr. Mark Steven Freedman, Professor of Medicine (Neurology)  
University of Ottawa, The Ottawa Hospital-General Campus  
501 Smyth Road  
Ottawa, Ont. K1H 8L6 (Canada)  
Tel. +1 613 737 8917, Fax +1 613 737 8857, E-Mail [mfreedman@ottawahospital.on.ca](mailto:mfreedman@ottawahospital.on.ca)

Another formulation of IFN $\beta$ -1a (Rebif<sup>®</sup>) was approved in 1998 in Europe and Canada and in 2002 in the United States; it is indicated for the treatment of RRMS to decrease the frequency of relapses and to delay the accumulation of physical disability. Mitoxantrone (Novantrone<sup>®</sup>), which has both immune-suppressive and immune-modulating effects, was approved by the FDA in 2000 for secondary progressive multiple sclerosis and worsening relapsing forms of multiple sclerosis (MS); however, it is not recommended as first-line therapy. A monoclonal antibody against  $\alpha$ -4 integrin, natalizumab (Tysabri<sup>®</sup>), was initially approved by the FDA in November 2004, but was withdrawn from the market in February 2005 due to safety concerns; it was re-introduced in June 2006 as monotherapy for the treatment of patients with relapsing forms of MS, to delay the accumulation of disability and to reduce the frequency of relapses.

Selecting a specific DMT for an individual patient involves a benefit-risk analysis, along with consideration of patient compliance and preference. The various therapies have different safety and tolerability profiles and require different frequencies and routes of administration. Pivotal clinical studies differ in their duration, retention rate, sample sizes, study populations, and placebo-effect sizes. In some cases the studies used differing primary endpoints, therefore making cross-comparisons difficult and not all that meaningful. The differing behaviors of the placebo groups in particular across the studies and over time show that, although patient populations are chosen on the basis of similar inclusion/exclusion characteristics and are expected to have the same risk of further MS-related activity, they do not. Since treatment effects are always measured relative to the individual study's placebo arm, it is misleading to simply compare these relative risk reductions (RRRs) across studies. The best way to avoid any bias is to perform direct head-to-head comparisons of therapies in the same population. Although others are forthcoming, only two such trials have so far been performed: the EVIDENCE (Evidence for Interferon Dose-Response: European North American Comparative Efficacy) [1, 2] and INCOMIN (Independent Comparison of Interferons) [3] trials.

Evidence-based medicine (EBM) is a systematic approach to clinical problem solving. Investigators using the EBM approach describe treatment effects arising from clinical trials in terms of neutral measures that tend to reduce reported relative differences to more modest and realistic effects that may be more comparable. Differences across clinical trials are reduced, though not eliminated, by looking at the overall therapeutic gain, or absolute risk

reduction, of a treatment. Therapeutic gain is simply the absolute difference between the outcome of the placebo and treatment groups, and is equivalent to the absolute risk reduction (ARR). When the apparent risk of an event is low, RRRs tend to over-represent treatment effects, whereas ARR is a better reflection of drug benefit. Because in MS clinical outcomes (relapses or progressions) are usually low in number, ARR should be used for comparisons. In terms of magnetic resonance imaging (MRI) parameters, the lesional change over time tends to be much higher and relative values become meaningful, though ARR is still more realistic to use for measures such as 'proportion free of enhancing lesions' that would often represent a very low number. We undertook a review of all the available data for the DMTs and compared the outcomes of the trials by looking at ARR, therapeutic gain, or number needed to treat (NNT), which is the measure favored in pharmacoeconomics. Such comparisons are possible in RRMS because of the consistent nature of the data captured for both clinical and MRI endpoints, despite differences in primary endpoints chosen across trials of DMTs.

## Materials and Methods

### Study Inclusion

In order to be included in this review, the study had to meet the following inclusion criteria: double-blind, placebo-controlled, randomized, multicenter trial; results published in a peer-reviewed journal or data available in FDA Summary Basis for Approval; patient population of clinically-definite RRMS; patient age  $\geq 18$  years; standard clinical or MRI efficacy endpoints available for analysis, and sample size  $>30$  patients. Ideally, the study should have been at least 2 years in duration, but shorter studies were considered if longer ones were not available.

On initial review, significant methodological issues with some peer-reviewed, published results were identified. For instance, in one clinical trial, data for 43% of the original intent to treat (ITT) population were not in the published results but were available in the FDA's *Summary Basis for Approval* for the therapy in question [4]. In another, due to a failure of randomization, the baseline characteristics of the active and placebo groups were not matched [5]. There were significant differences in how this failure was addressed statistically by the authors and by the FDA. Because of these types of issues, only trials that were reviewed independently by the FDA were considered for analysis. In those cases where the published data were not available or conflicted with the FDA's analyses, the FDA's analyses were used.

### Clinical Efficacy Endpoints

Three different categories of clinical efficacy endpoints were examined. The first category pertained to relapse-related endpoints (relapse was defined as acute worsening of status of at least 24 h duration, where worsening of status was indicated by a change in the EDSS score of at least 1 point). The various possible relapse-

related endpoints were: time to first relapse; proportion of patients who were relapse-free at 1 year; proportion of patients who were relapse-free at 2 years; number of relapses; annualized relapse rate, and percent reduction in the annualized relapse rate.

The second category of clinical efficacy endpoints was progression/disability-related, with progression defined as a change in the EDSS score of at least 1 point that lasted at least 3 months. Possible progression-related endpoints included time to confirmed progression in disability, proportion of patients with no progression in confirmed disability at 1 year, proportion of patients with no progression in confirmed disability at 2 years, and percent reduction in confirmed disability at 2 years.

The third category consisted of MRI endpoints. Possible endpoints included the number of new gadolinium (Gd)-enhancing lesions, percent reduction in new Gd-enhancing lesions, the number of new T2 lesions, percent reduction in new T2 lesions, percent increase or reduction in T2 burden of disease (i.e. total T2 lesion area or total T2 lesion volume), or the proportion of patients free of Gd-enhancing lesions at 1 year.

Endpoints for which there were insufficient comparative data among the accepted studies were not included in the analyses. Specifically, for many of the continuous variables the standard error of measurement was not publicly available, making cross comparisons difficult using accepted statistical techniques.

#### *Statistical Analysis*

The statistical analysis was performed by PROMETRIKA, an independent contract research organization. Statistical analyses were performed using Review Manager software (version 4.2) and SAS® for Windows (version 8.2) [6, 7].

Risk differences (RDs, another expression used for ARR) and corresponding 95% confidence intervals (CIs) were calculated by study for each binary outcome. Differences between the active and placebo groups allowed for the comparison of active treatments across studies based on the concept of ARR and therapeutic gain. Relative risks (RR) and their 95% CIs were also calculated for each binary outcome for the purpose of sensitivity analyses and for consistency with previously published literature in this area of research [8]. For continuous outcomes, rate differences of the active and placebo means were calculated, as were 95% CIs for the rate differences within each study. Standard deviations were not always available for calculation of the CIs; however, when p values for treatment mean comparisons were presented, the standard error for the difference of the means was derived using percentiles of the t-distribution.

Because the goal of the analysis was to make comparisons among the various MS treatments with respect to the efficacy endpoints, the calculation of an overall treatment effect across studies was not performed. In studies where more than one dosage was used, only data for the highest dosage used were analyzed, in order to demonstrate maximal clinical effect.

## **Results**

### *Studies and Endpoints Selected*

Six studies involving the 5 FDA-approved first-line DMTs met the inclusion criteria and were selected for

further analysis (table 1). The EVIDENCE [1, 2], INCOMIN [3] and OWIMS (Once Weekly Interferon for MS) studies [9] were considered but not selected for analysis because they did not meet all of the inclusion criteria. Two studies were included for Copaxone and 1 study each for the 4 other DMTs. The data used from the PRISMS (Prevention of Relapses and Disability by Interferon  $\beta$ -1a Subcutaneously in MS) study [10] was for the higher dose of Rebif (44  $\mu$ g) because the lower dose (22  $\mu$ g) was not FDA-approved at the time of the analysis and the protocol dictated that only the highest dose be analyzed for studies with multiple dosages. Similarly, only data for the higher dose of Betaseron (8 MIU s.c.) was used.

Data from the included studies were available for the following efficacy endpoints: proportion of patients who were relapse-free at 1 and 2 years; annualized relapse rate at 2 years; proportion of patients who were progression-free at 1 and 2 years, and proportion of patients who were free of Gd-enhancing lesions at 1 year or 9 months. Not all of these endpoints were available from all of the 6 included studies.

### *Relapse Endpoints*

#### *Proportion of Patients Who Were Relapse-Free at 1 Year or 9 Months*

RDs (proportion of active patients minus proportion of placebo patients) and their corresponding 95% CIs for the proportion of patients who were progression-free at 1 year are presented in figure 1a. The RDs for both Avonex and Copaxone (based on 9-month data) were not significantly different from zero, i.e. the proportion of patients who were relapse-free at 1 year was not significantly greater for Avonex or Copaxone than placebo. The RDs for Rebif and Tysabri were significantly greater than zero, i.e. the proportion of patients who were relapse-free at 1 year was significantly greater than placebo for both agents. The ARR, therapeutic gain, and NNT (1/ARR) for the different agents are shown in table 2. The values for NNT range from 4.3 for Rebif and Tysabri to 16.7 for Copaxone.

The relative risks (RRs, proportion of active patients divided by proportion of placebo patients) and their corresponding 95% CIs for the proportion of patients who were relapse-free at 1 year are presented in figure 1b. The RRs for both Avonex and Copaxone were not significantly different from 1 (i.e. the proportion of patients who were relapse-free at 1 year was not significantly greater for Avonex or Copaxone than placebo). The RRs for Rebif and Tysabri were significantly greater than 1 (i.e. the proportion of patients who were relapse-free at 1 year were

**Table 1.** Characteristics of included studies

Study and drug [reference]	Location (centers)	Enrollment period	Study duration months	Treatments (number randomized)	Dosing schedule	Patient baseline characteristics
<b>Interferons</b>						
MSCRG Avonex [4, 21]	USA (4)	1990–1993	24	30 µg interferon beta-1a (158) placebo (143)	IM, weekly	age 18–55; definite RRMS; EDSS 1–3.5; disease duration ≥1 year; at least 2 relapses in the 3 years prior; no exacerbations for at least 2 months prior
PRISMS Rebif [10, 22, 23]	Canada, Europe (22)	1994–1995	24	22 µg interferon beta-1a (189) 44 µg interferon beta-1a (184) placebo (187)	SC, 3 times weekly	age not reported; clinical or laboratory-supported definite RRMS; EDSS 0–5.0; disease duration ≥1 year; at least 2 relapses in the 2 years prior
IFNB Betaseron [24–26]	USA, Canada (11)	1988–1990	24 (plus 42-month extension)	0.05 mg interferon beta-1b (125) 0.25 mg interferon beta-1b (124) placebo (123)	SC, every other day	age 18–50; clinical or laboratory-supported definite RRMS; EDSS 0–5.5; disease duration >1 year; at least 2 relapses in the 2 years prior; no exacerbations for at least 1 month prior
<b>Glatiramer acetate</b>						
Johnson et al. Copaxone [16, 17]	USA (11)	1991–1992	24 (plus 11-month extension)	20 mg glatiramer acetate (125) placebo (126)	SC, daily	age not reported; well-characterized RRMS; EDSS 0–6.0; disease duration not specified; at least 2 relapses in the 2 years prior
Comi et al. Copaxone [27]	Canada, Europe (29)	1997	9	20 mg glatiramer acetate (119) placebo (120)	SC, daily	age 18–50; diagnosis of RRMS, including ≥1 enhancing brain lesion; EDSS 0–5.0; disease duration ≥1 year; at least 1 relapse in the 2 years prior; no exacerbations in previous 30 days
<b>Immunomodulator</b>						
Polman et al. Tysabri [13, 20]	Europe, N. America, Australia, New Zealand (99)	2001–2004	28	300 mg natalizumab (627) placebo (315)	IV infusion, every 4 weeks	age 18–50; diagnosis of RRMS, including brain lesions consistent with MS; EDSS 0–5.0; disease duration not specified; at least 1 relapse in the past year
EDSS = Expanded disability status scale; IM = intramuscular; IV = intravenous; MIU = million international units; SC = subcutaneous.						

significantly greater than placebo for both Rebif and Tysabri). The RR results are consistent with those from the RD analyses and indicate that both Rebif and Tysabri are significantly better than placebo in preventing relapses during the first year of treatment.

#### Proportion of Patients Who Were Relapse-Free at 2 Years

RDs and their corresponding 95% CIs for the proportion of patients who were relapse-free after 2 years of treatment are presented in figure 2a. Relapse-free data at 2 years were available for Betaseron, to compare to the relapse-free data at 1 year. Relapse-free data at 2 years were not available for all Avonex patients, because the study was ended prematurely. It should be noted that the 9-month and 2-year Copaxone data were from 2 separate studies. Copaxone was not significantly different from placebo; the other DMTs were significantly better than

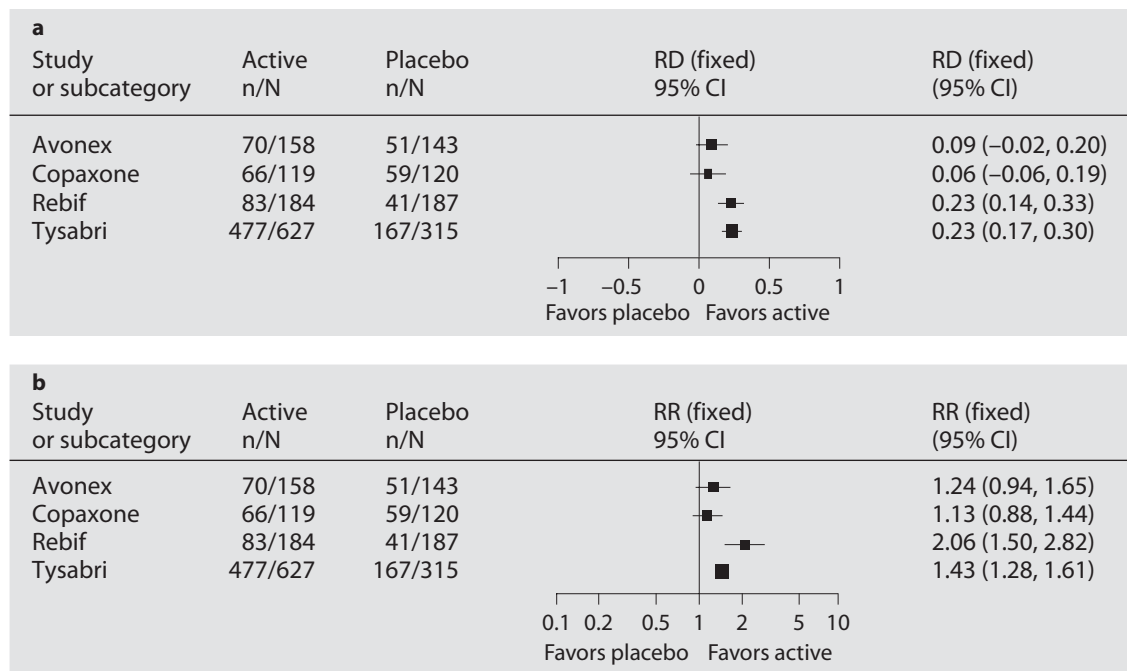
placebo and had RDs of similar magnitude. The values for NNT for the different agents ranged from 3.8 for Tysabri to 14.3 for Copaxone (table 2).

Rrs and 95% CIs for the proportion of patients who were relapse-free after 2 years of treatment are presented in figure 2b. Again, Copaxone was not significantly different from placebo; the other DMTs were significantly better than placebo. Overall, the 1-year and 2-year data provide evidence that Rebif, Betaseron and Tysabri are more effective than placebo in maintaining patients with RRMS free of attacks.

#### Annualized Relapse Rates at 2 Years

Differences between the mean annualized relapse rates of the active and placebo groups in each study and 95% CIs for the differences in means are presented in figure 3. All the DMTs were significantly better than placebo. Overall, there did not appear to be a clear advantage





**Fig. 1.** Proportion of patients who were relapse-free at 1 year: Representation of the efficacy of treatment with the various agents relative to placebo for different endpoints. Larger positive values indicate better efficacy relative to placebo. Relapse was defined as acute worsening of status of at least 24 h duration, where worsening of status was indicated by a change in the Expanded Disability Scale score of at least 1 point. **a** Risk difference (RD), which is equivalent to absolute risk reduction (ARR). **b** Relative risk reduction (RRR).

of one treatment over the others in terms of RD, except for Avonex, which had a lower annualized relapse rate and reached only borderline significance. This analysis included all patients and all of their time in the study; this was possible with the annualized rate.

#### Progression/Disability Endpoints

##### Proportion of Patients Who Were Progression-Free at 1 Year

Progression-free data at 1 year were available only for Avonex. Following 1 year of treatment, 83% of Avonex patients and 73% of placebo patients were relapse-free, which corresponds to an RD of 0.10 (95% CI 0.00–0.19) and a RR of 1.13 (95% CI 1.00–1.27). Thus, Avonex was not significantly better than placebo in preventing disease progression at 1 year.

##### Proportion of Patients Who Were Progression-Free at 2 Years

Progression-free data at 2 years were available for 5 DMTs. Of the 5 trials, only the Avonex and Tysabri trials had progression-related endpoints as a primary efficacy outcome and were appropriately powered to show a treat-

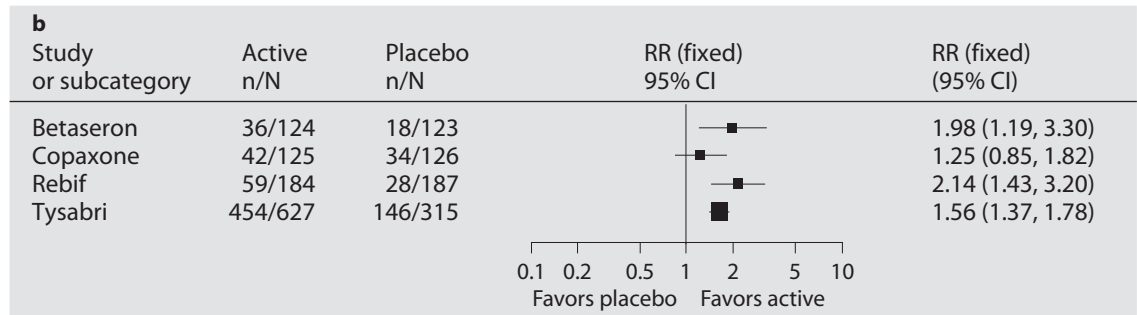
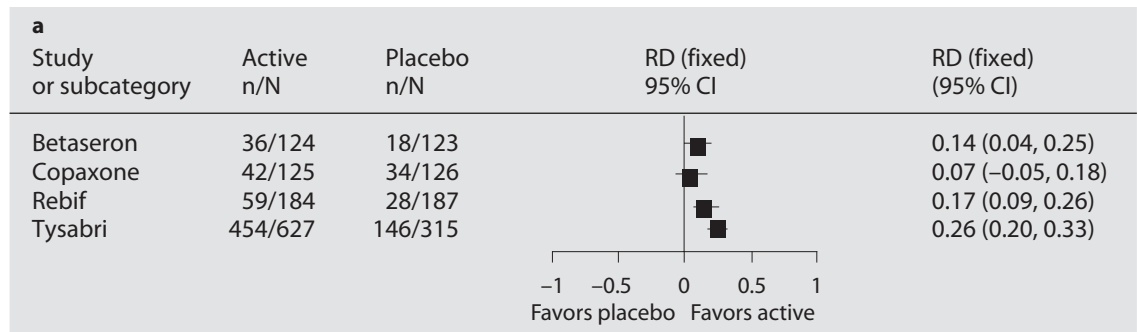
ment effect. The other trials were powered to show efficacy in terms of relapse-related endpoints; progression-related endpoints were assessed as secondary outcomes.

RDs and their corresponding 95% CIs for the proportion of patients who were progression-free at 2 years are presented in figure 4a. All of the treatments had RDs ranging from 0.10 to 0.13, except Copaxone, which had an RD of 0.03. The corresponding values for NNT ranged from 7.7 to 33.3 (table 2). Results from analyses of the RRs, shown in figure 4b, paralleled those of the RDs. All treatments appeared to provide similar benefit in preventing progressions over 2 years of treatment, except for Copaxone and Betaseron, which were not significantly different from placebo.

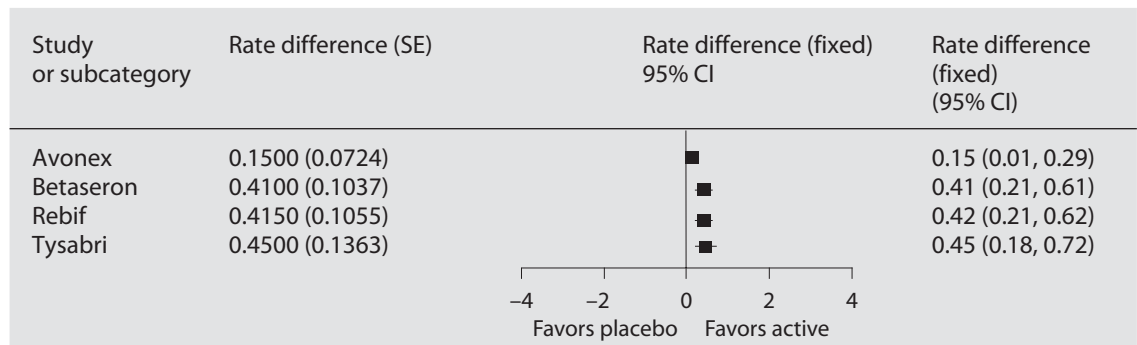
#### MRI Endpoint

##### Proportion of Patients Who Were Free of Gd-Enhancing Lesions at 1 Year or 9 Months

RDs and their corresponding 95% CIs for the proportion of patients who were free of Gd-enhancing lesions at 1 year are presented in figure 5a. In most cases (except Tysabri), only a subset of the ITT population underwent Gd-enhanced MRI scanning. Also, because the Betaser-



**Fig. 2.** Proportion of patients who were relapse-free at 2 years. Representation of the efficacy of treatment with the various agents relative to placebo for different endpoints. Larger positive values indicate better efficacy relative to placebo. Relapse was defined as acute worsening of status of at least 24 h duration, where worsening of status was indicated by a change in the Expanded Disability Status Scale score of at least 1 point. **a** Risk difference (RD), which is equivalent to absolute risk reduction (ARR). **b** Relative risk reduction (RRR).



**Fig. 3.** Annualized relapse rate at 2 years. Representation of the efficacy of treatment with the various agents relative to placebo for different endpoints. Larger positive values indicate better efficacy relative to placebo. Relapse was defined as acute worsening of status of at least 24 h duration, where worsening of status was indicated by a change in the Expanded Disability Status Scale score of at least 1 point. Figure shows rate difference (RD), which is equivalent to absolute risk reduction (ARR).

on study was performed prior to the widespread use of Gd-enhanced MRI, data for Betaseron were unavailable. Only Copaxone did not show a significant benefit for this endpoint; however, its onset of action may be delayed for up to 6 months, and this was just a 9 month study. More-

over, the Copaxone study was different in that the study population needed to have  $\geq 1$  enhancing lesion as an inclusion criterion. The values for NNT ranged from 3.2 for Rebif to  $\infty$  for Copaxone (table 2).

**Table 2.** Absolute risk reduction, therapeutic gain and number needed to treat for the different agents

	AAR, therapeutic gain (95% CI)	NNT (95% CI)
Proportion of patients who were relapse-free at 1 year		
Avonex	0.09 (-0.02, 0.20)	11.1 (5.0, ∞)
Copaxone	0.06 (-0.06, 0.19)	16.7 (5.3, ∞)
Rebif	0.23 (0.14, 0.33)	4.3 (3.0, 7.1)
Tysabri	0.23 (0.17, 0.30)	4.3 (3.3, 5.9)
Proportion of patients who were relapse-free at 2 years		
Betaseron	0.14 (0.04, 0.25)	7.1 (4.0, 25.0)
Copaxone	0.07 (-0.05, 0.18)	14.3 (5.6, ∞)
Rebif	0.17 (0.09, 0.26)	5.9 (3.8, 11.1)
Tysabri	0.26 (0.20, 0.33)	3.8 (3.0, 5.0)
Proportion of patients who were progression-free at 2 years		
Avonex	0.13 (0.03, 0.23)	7.7 (4.3, 33.3)
Betaseron	0.10 (-0.01, 0.21)	10.0 (4.8, ∞)
Copaxone	0.03 (-0.07, 0.13)	33.3 (7.7, ∞)
Rebif	0.11 (0.01, 0.20)	9.1 (5.0, 100)
Tysabri	0.12 (0.06, 0.18)	8.3 (5.6, 16.7)
Proportion of patients who were free of Gd-enhancing lesions at 1 year		
Avonex	0.12 (0.01, 0.24)	8.33 (4.2, 100)
Copaxone	-0.03 (-0.08, 0.02)	∞ (50.0, ∞)
Rebif	0.31 (0.17, 0.44)	3.2 (2.3, 5.9)
Tysabri	0.28 (0.23, 0.33)	3.6 (3.0, 4.3)

AAR = Absolute risk reduction; Gd = gadolinium; NNT = number needed to treat.

RRs and 95% CIs for the proportion of patients who were free of new Gd-enhancing lesions at 1 year are presented in figure 5b. Copaxone is indistinguishable from placebo, but this study differed from the other treatments due to differences at baseline. Of the DMTs other than Copaxone, Rebif appears to be better at preventing new Gd-enhancing lesions than Tysabri or Avonex. These results suggest there may be a benefit of Rebif over these other treatments in the prevention of new Gd-enhancing lesions in patients with RRMS at 1 year.

## Discussion

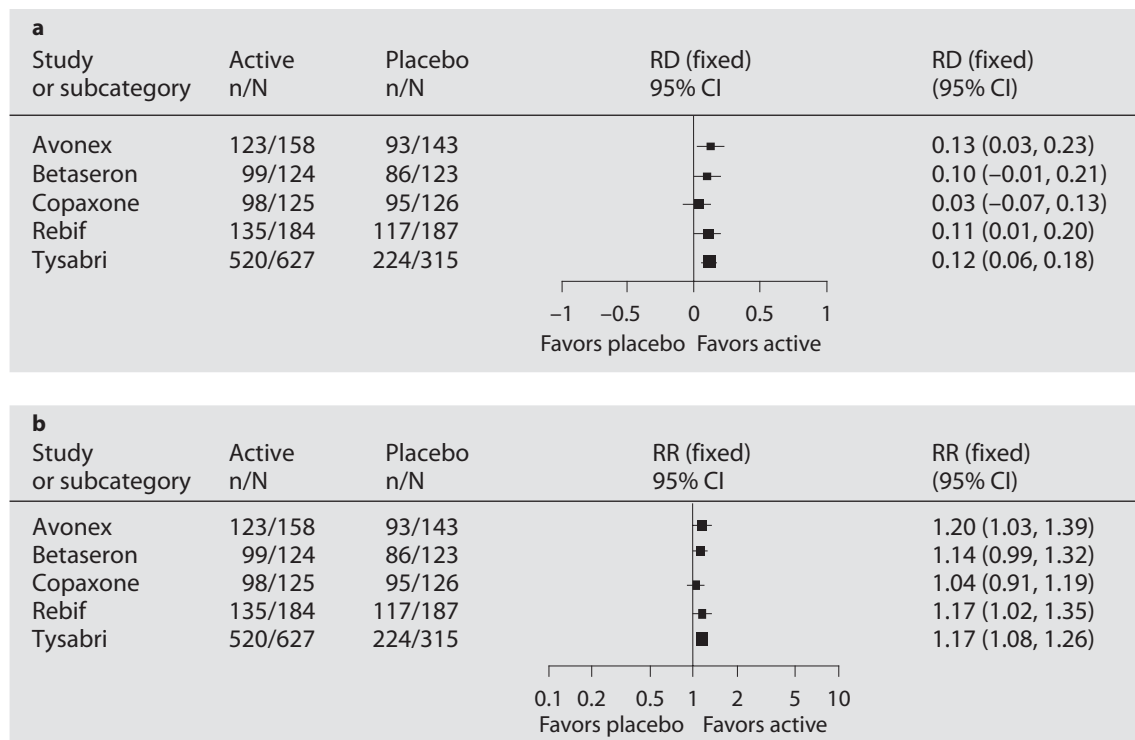
Though we now have 5 approved products for the treatment of RRMS, there are many differences among the pivotal phase III studies that preclude simply comparing their efficacy outcomes, which were all calculated relative to different placebo groups. We used an evidence-based approach to attempt to reduce these outcomes to a

more common denominator, such as RD or NNT, in order to better compare efficacy.

In general, there was little difference among the therapies in terms of clinical endpoints such as their effect on relapse or progression; annualized relapse rates were less sensitive to differences than the proportions of patients who were relapse-free. All the IFN treatments and Tysabri showed similar efficacy in terms of increasing the percentage of patients who were progression-free at the end of 2 years. Copaxone failed to increase this percentage, possibly in part due to the study being statistically underpowered to be able to demonstrate this, similar to the pivotal Betaseron trial. Trials with Avonex, Rebif and Tysabri all were structured to be able to show an effect on progression as a primary or secondary endpoint. Coupled with underpowering is the phenomenon that only a small percentage of RRMS patients would be expected to have a sustained progression within 2 years. It is not surprising that the most sensitive efficacy endpoint appeared to be the prevention of Gd-enhancing lesions (RRRs ranging from 0.50 to 3.88), which is, in part, because of greater event rates compared with either relapse or progression. However, there remains great debate on the reliability and validity of standard MRI measures as surrogate biomarkers in MS [11, 12].

The proportion of patients who were relapse-free at 1 year and the proportion of patients who were free of Gd-enhancing lesions at 1 year showed the clearest differentiation among the various DMTs. Though 1-year data may not be long enough for therapies to reach maximal efficacy, we felt it was important to explore these in the context of early treatment response. Some studies suggest that it may be possible to demonstrate a treatment effect in 1 year in order to decide whether patients should complete the course or switch to an alternative therapy. Thus it was important to see if agents were capable of generating a treatment response in 1 year and try to compare them. In this regard, both Tysabri and Rebif appeared to provide a greater 1-year benefit than Avonex.

In looking at all the efficacy information from pivotal, placebo-controlled, independently-reviewed clinical trials for the currently-available DMTs, a general picture emerges. Of the available agents, Tysabri and high-dose, high-frequency IFNs (Betaseron and Rebif) appear to be significantly more effective than placebo. Keeping in mind that study design and execution have evolved considerably over the years, and notwithstanding the differences in sample sizes and patient characteristics across the trials, Rebif (PRISMS) [10] and Tysabri [AFFIRM (Natalizumab Safety and Efficacy in Relapsing Remitting



**Fig. 4.** Proportion of patients who were progression-free at 2 years. Representation of the efficacy of treatment with the various agents relative to placebo for different endpoints, with larger positive values indicating better efficacy relative to placebo. Progression was defined as a change in the Expanded Disability Status Scale score of at least 1 point that lasted at least 3 months. **a** Risk difference (RD), which is equivalent to absolute risk reduction (ARR). **b** Relative risk reduction (RRR).

Multiple Sclerosis) study] [13] stand out as the only agents that significantly affect all the measured efficacy outcomes of relapse, progression and MRI. Furthermore, our comparison shows that Rebif and Tysabri have a similar magnitude of effect on most of these efficacy parameters. Both Avonex and Copaxone show, in general, less robust effects, with no definite benefit over placebo for many of the endpoints studied.

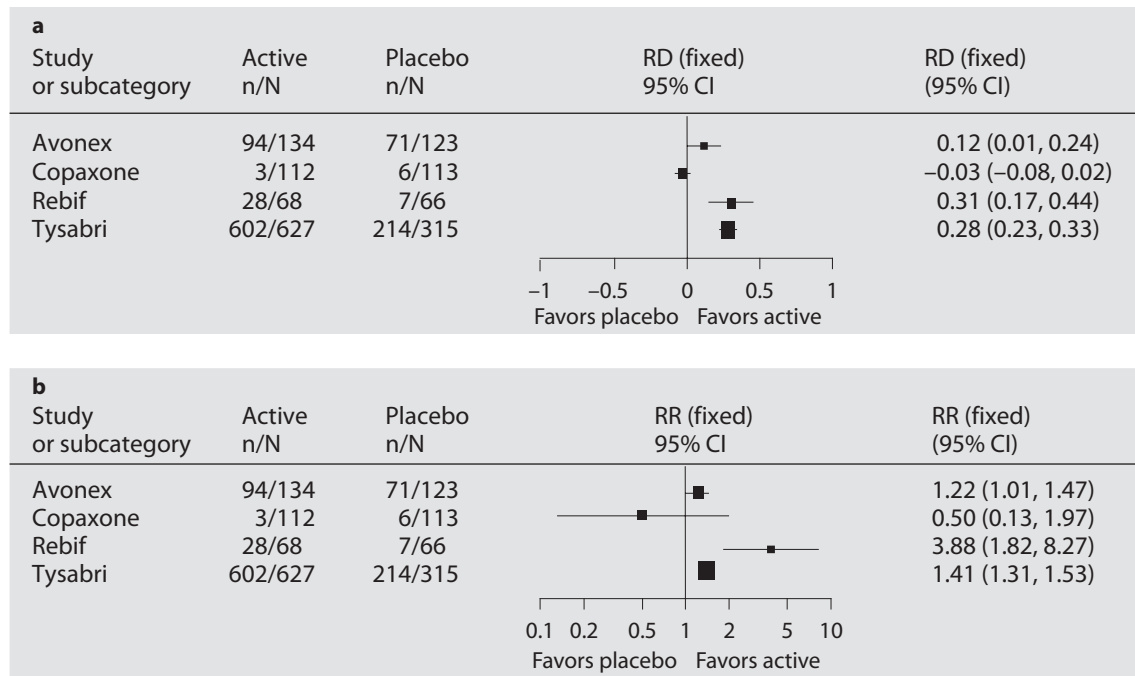
It is important to keep in mind the many differences in study design for the various DMTs that could lead to disparity in clinical results. For instance, the AFFIRM trial (Tysabri) used the McDonald criteria [14] for MS diagnosis versus the Poser criteria [15], which was used in all the prior studies. Thus, the AFFIRM trial included a population of patients that were younger, less disabled and with a shorter duration of disease, all factors that might impact treatment efficacy. One of the most important differences among the trials is the natural history of the placebo group. In trials like the AFFIRM or the pivotal Copaxone study, the reduction in relapse rate at 2

years relative to baseline (placebo effect) was 51% [13] and 47% [16, 17], respectively. Such patients may have a very different natural history compared with those in the other studies like PRISMS (Rebif), where the natural risk reduction in the placebo group after 2 years was only 17% [10].

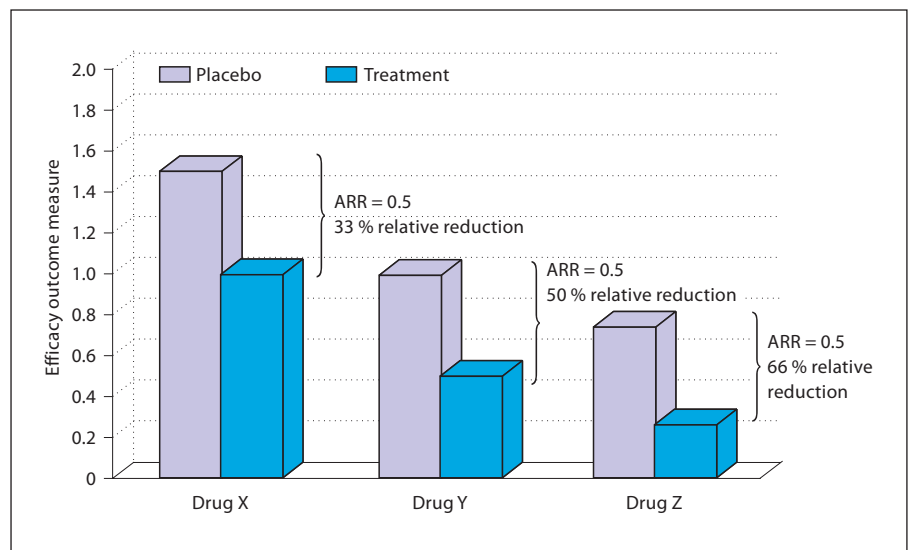
The results of this review can be compared with the published results of some other clinical studies in patients with RRMS. In the OWIMS study, Rebif administered weekly for 24 or 48 weeks resulted in a significant MRI benefit but not a clinical relapse effect relative to placebo [9, 18]. However, when the same dose of Rebif was given to patients after their first demyelinating episode in the ETOMS (Early Treatment of MS) study [19], there was a significant effect on both relapse rate and MRI. These studies suggest there may be differences in treatment response to DMT depending on the stage of disease or its duration.

In the 2 available head-to-head studies (EVIDENCE [1, 2] and INCOMIN [3]) comparing frequent, high-dose





**Fig. 5.** Proportion of patients who were free of Gd-enhancing lesions at 1 year. Representation of the efficacy of treatment with the various agents relative to placebo for different endpoints. Larger positive values indicate better efficacy relative to placebo. **a** Risk difference (RD), which is equivalent to absolute risk reduction (ARR). **b** Relative risk reduction (RRR). Note that the data for Copaxone is for 9 months.



**Fig. 6.** Absolute versus relative risk reduction for 3 hypothetical drugs, X, Y and Z. The absolute risk reduction (ARR) is the same for all 3 drugs (0.5). Drug Z does not really demonstrate twice the efficacy of drug X.

IFN $\beta$  treatment with once weekly Avonex, both showed that high-dose IFNs (Rebif and Betaseron) are superior compared to low-dose IFN (Avonex) in clinical efficacy at 1 or 2 years of treatment.

The difference between the use of ARR and RRR is demonstrated by figure 6. How the placebo group fares in a particular study gives an indication of the risk of having a particular event. If 2 populations have placebo groups

that demonstrate similar risks, then it may be possible to compare the RRRs conferred by treatment. However, if the placebo groups differ greatly in their risk of an event, then RRRs are not comparable and it is better to use ARR or therapeutic gain as a comparator. Take, for example, 3 hypothetical drugs, X, Y and Z (fig. 6). The placebo rate of an event differs greatly among the 3 agents such that, if one only looks at RRR, it would appear that drug Z is nearly twice as efficacious as drug X. This is clearly misleading as all show an ARR of 0.5 or an NNT of 2. How drug X would have done in a population such as that tested with drug Z, or vice versa, is unknown. This challenge is relevant to the current literature on MS. The AFFIRM study [13] found that there was a 68% relative reduction in the rate of relapse with Tysabri. The PRISMS study [10] found that the relapse rate reduction for Rebif (44 µg) was 32%. However, comparisons of these RRRs may be misleading as the ARR for Tysabri is 0.26 (95% CI 0.20–0.33; NNT = 3.8) compared with an ARR for Rebif of 0.17 (95% CI 0.09–0.26; NNT = 5.9) over the same 2-year period (table 2). Further, the confidence intervals for these ARR values overlapped, providing descriptive evidence that the NNT estimates may not be different from each other.

As the treatment paradigm of MS evolves and newer agents become available, making treatment decisions and providing guidance for patients will become more challenging for clinicians. There are several ongoing studies comparing directly the efficacy of high-dose, high-frequency IFNβs against Copaxone, which will shed light on the relative efficacies of these agents. Ideally, one would want to compare their effects relative to a common placebo group; however, due to ethical concerns, it is extremely difficult to carry out such types of head-to-head, double-blinded, placebo-controlled studies in countries where DMTs are available. A way of performing a head-to-head study without the use of a placebo group is to carry out a ‘double dummy’ study where patients literally inject themselves as if they were taking both medicines, but are in reality only taking one. This is the method used in the ongoing Combi-Rx study that is looking at three groups: Avonex only, Copaxone only and both. In the absence of these types of studies, one can look at the difference in proportions (RD) and RR to provide some sense of relative efficacy in order to guide prescribing clinicians. Still, there is no substitute for well designed, head-to-head, class I, comparative studies (such as the EVIDENCE [1, 2] and INCOMIN [3] trials), which would resolve all of the cross-study comparison issues.

Using an EBM approach to study results can yield important information that will assist both physicians and patients in making treatment decisions. This is because the way in which trial data are conveyed to patients can influence their ultimate decision and enthusiasm for one product over another. For example, patients could be told that there were 68% fewer attacks with Tysabri compared to a non-treated group [13], and that there were 32% fewer attacks with Rebif (44 µg) compared to a non-treated group [10]. Although the numbers are not directly compared, simply stating them one after another would suggest a stronger treatment effect for Tysabri than Rebif. Using the EBM approach, the patient could be told that the number of patients who needed to be treated to prevent 1 relapse is 3.8 for Tysabri, compared to 5.9 for Rebif. Thus, the efficacies of Tysabri and Rebif in preventing relapses are very similar based on the NNT, and choosing between the agents may require a consideration of safety issues. In particular, Tysabri has a black box warning about the risk of progressive multifocal leukoencephalopathy [20] and thus poses additional risks compared to Rebif.

Finally, data from placebo-controlled clinical trials can be used to predict what should be observed in head-to-head studies. This analysis is in progress and the calculations will be compared with the results observed in the EVIDENCE [1, 2] and INCOMIN [3] clinical trials.

### Acknowledgements

We would like to thank Carol Lewis, PhD for her editorial assistance in the preparation of the manuscript. EMD Serono, Inc. provided financial support for the preparation of the manuscript.

### References

- 1 Panitch H, Goodin D, Francis G, Chang P, Coyle P, O'Connor P, Li D, Weinshenker B: Benefits of high-dose, high-frequency interferon beta-1a in relapsing-remitting multiple sclerosis are sustained to 16 months: final comparative results of the EVIDENCE trial. *J Neurol Sci* 2005;239:67–74.
- 2 Panitch H, Goodin DS, Francis G, Chang P, Coyle PK, O'Connor P, Monaghan E, Li D, Weinshenker B: Randomized, comparative study of interferon beta-1a treatment regimens in MS: the EVIDENCE Trial. *Neurology* 2002;59:1496–1506.

- 3 Barbero P, Bergui M, Versino E, Ricci A, Zhong JJ, Ferrero B, Clerico M, Pipieri A, Verdun E, Giordano L, Durelli L: Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis (INCOMIN Trial) II: analysis of MRI responses to treatment and correlation with Nab. *Mult Scler* 2006;12:72–76.
- 4 Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, Fischer JS, Goodkin DE, Granger CV, Simon JH, Alam JJ, Bartoszak DM, Bourdette DN, Braiman J, Brownschieldle CM, Coats ME, Cohan SL, Dougherty DS, Kinkel RP, Mass MK, Munschauer FE 3rd, Priore RL, Pullicino PM, Scherokman BJ, Whitham RH; The Multiple Sclerosis Collaborative Research Group (MSCRG): Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Ann Neurol* 1996;39:285–294.
- 5 Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP, Myers LW, Panitch HS, Rose JW, Schiffer RB; The Copolymer 1 Multiple Sclerosis Study Group: 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;45:1268–1276.
- 6 RevMan 4.2 User Guide. Oxford, The Cochrane Collaboration, 2004.
- 7 SAS/STAT User's Guide, Version 8. Cary, SAS Institute, Inc., 1999.
- 8 Filippini G, Munari L, Incorvaia B, Ebers GC, Polman C, D'Amico R, Rice GP: Interferons in relapsing remitting multiple sclerosis: a systematic review. *Lancet* 2003;361:545–552.
- 9 The Once Weekly Interferon for MS Study Group: Evidence of interferon beta-1a dose response in relapsing-remitting MS: the OWIMS Study. *Neurology* 1999;53:679–686.
- 10 PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group: Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. *Lancet* 1998;352:1498–1504.
- 11 Miller DH: Biomarkers and surrogate outcomes in neurodegenerative disease: lessons from multiple sclerosis. *NeuroRx* 2004;1:284–294.
- 12 Schreiber K, Sorensen PS, Koch-Henriksen N, Wagner A, Blinkenberg M, Svarer C, Petersen HC: Correlations of brain MRI parameters to disability in multiple sclerosis. *Acta Neurol Scand* 2001;104:24–30.
- 13 Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L, Miller DH, Phillips JT, Lublin FD, Giovannoni G, Wajgt A, Toal M, Lynn F, Panzara MA, Sandrock AW: A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med* 2006;354:899–910.
- 14 McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, McFarland HF, Paty DW, Polman CH, Reingold SC, Sandberg-Wollheim M, Sibley W, Thompson A, van den Noort S, Weinshenker BY, Wolinsky JS: Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001;50:121–127.
- 15 Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, Johnson KP, Sibley WA, Silberberg DH, Tourtellotte WW: New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983;13:227–231.
- 16 Copaxone package insert. Kansas City, Teva Neuroscience, Inc., 2004.
- 17 Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP, Myers LW, Panitch HS, Rose JW, Schiffer RB, Vollmer T, Weiner LP, Wolinsky JS; Copolymer 1 Multiple Sclerosis Study Group: Extended use of glatiramer acetate (Copaxone) is well tolerated and maintains its clinical effect on multiple sclerosis relapse rate and degree of disability. *Neurology* 1998;50:701–708.
- 18 Freedman MS, Francis GS, Sanders EA, Rice GP, O'Connor P, Comi G, Duquette P, Metz L, Murray TJ, Bouchard JP, Abramsky O, Pelletier J, O'Brien F: Randomized study of once-weekly interferon beta-1a therapy in relapsing multiple sclerosis: three-year data from the OWIMS study. *Mult Scler* 2005;11:41–45.
- 19 Filippi M, Rovaris M, Inglese M, Barkhof F, De Stefano N, Smith S, Comi G: Interferon beta-1a for brain tissue loss in patients at presentation with syndromes suggestive of multiple sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet* 2004;364:1489–1496.
- 20 Tysabri package insert. Cambridge, Biogen Idec, Inc., 2006.
- 21 Avonex Summary Basis for Approval. Cambridge, Biogen Index Inc., <http://www.fda.gov/cder/biologics/sba/ifnbbio051796s.pdf>.
- 22 Rebif Summary Basis for Approval – Statistical Review. Rockland, Serono, Inc., Feb 18, 1999.
- 23 Li DK, Paty DW: Magnetic resonance imaging results of the PRISMS trial: a randomized, double-blind, placebo-controlled study of interferon-beta1a in relapsing-remitting multiple sclerosis. Prevention of relapses and disability by interferon-beta1a subcutaneously in multiple sclerosis. *Ann Neurol* 1999;46:197–206.
- 24 Betaseron package insert. Emeryville, CA, Chiron Corp., 2003.
- 25 The IFNB 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;43:655–661.
- 26 The IFNB Multiple Sclerosis Study Group, The University of British Columbia MS/MRI Analysis Group: Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. *Neurology* 1995;45:1277–1285.
- 27 Comi G, Filippi M, Wolinsky JS; European/Canadian Glatiramer Acetate Study Group: European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging-measured disease activity and burden in patients with relapsing multiple sclerosis. *Ann Neurol* 2001;49:290–297.