

Technology Assessment



**Technology
Assessment Program**

**Agency for Healthcare
Research and Quality
540 Gaither Road
Rockville, Maryland 20850**

**The clinical effectiveness
and cost-effectiveness of
interferon-beta and
glatiramer acetate in the
management of
relapsing/remitting and
secondary-progressive
multiple sclerosis**

23rd October 2006

The clinical effectiveness and cost-effectiveness of interferon-beta and glatiramer acetate in the management of relapsing/remitting and secondary-progressive multiple sclerosis

Prepared by:

The School of Health and Related Research, The University of Sheffield, England

Paul Tappenden, MSc

Christopher McCabe, PhD

Emma Simpson, PhD

Jim Chilcott, MSc

Richard Nixon, PhD

Jason Madan, MSc

Acknowledgements

We would like to thank Professor Murray Brown, of the Department of Community Health and Epidemiology, Dalhousie University, Halifax, Nova Scotia, for providing health utility data from the Nova Scotia Multiple Sclerosis Integrated Database. We would also like to thank Dr Nicholas LaRocca of the National Multiple Sclerosis Society for granting the project team access to the Sonya Slifka database. We would like to acknowledge the considerable body of work undertaken by ABT Associates, Inc. in analyzing the Sonya Slifka database and Medicare claims data. We would also like to acknowledge the earlier input of the Cost Effectiveness of MS Therapies Working Group, which formed in 2001 during the National Institute for Clinical Excellence's appraisal of beta interferon and glatiramer acetate for the management of MS.

Contribution of authors

Paul Tappenden, Jim Chilcott and Christopher McCabe developed the ScHARR MS Cost-effectiveness Model in conjunction with the Cost Effectiveness of MS Therapies Working Group. Emma Simpson and Paul Tappenden undertook the review of clinical effectiveness. Paul Tappenden, Jim Chilcott and Christopher McCabe developed the cost-effectiveness model for use by the Centers for Medicare and Medicaid. Richard Nixon, Jason Madan, Paul Tappenden and Jim Chilcott undertook the Bayesian synthesis of RCT evidence using WinBUGS.

Disclaimer

This project was funded under Contract No. RJ108076 from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Conflicts of interest

Since the 2001 NICE Appraisal, authors of this report have provided advice concerning modeling the cost-effectiveness of treatments for MS to Biogen IDEC and Novartis. None of the treatments involved are within the scope of the AHRQ analysis. Advice is currently being given to the US Multiple Sclerosis Society.

4.10 Conclusions on clinical effectiveness	39
4.11 Summary of clinical effectiveness data used to populate the health economic model	41
5.0 Health economic methods	43
5.1 Introduction	43
5.2 Existing evidence on the health economics of IFN β and GA	43
5.3 Perspective of the analysis	44
5.4 Structure of health economic model	44
5.5 Health economic comparisons undertaken within the cost-effectiveness analysis	46
5.6 Data sources used within the ScHARR MS cost-effectiveness model	47
6.0 Health economic results	64
6.1 Introduction	64
6.2 Central estimates of cost-effectiveness	64
6.3 Simple sensitivity analysis	65
6.4 Probabilistic sensitivity analysis	74
7.0 Discussion	78
7.1 Discussion of clinical effectiveness results	78
7.2 Discussion of health economic results	81
7.3 Further considerations	84
7.4 Limitations of the health economic analysis	85
7.5 Areas for further research	87
8.0 Conclusions	89
8.1 Conclusions on the clinical effectiveness of the disease-modifying therapies for MS	89
8.2 Conclusions on the cost-effectiveness of the disease-modifying therapies for MS	89
8.3 Concluding remarks	91
Appendices	
Appendix 1 Expanded Disability Status Scale	92
Appendix 2 Worked example of progression-free survival hazards	93
Appendix 3 WinBUGS mixed treatment model code	94
Appendix 4 List of included and excluded studies	96
Appendix 5 List of model parameters	98
Appendix 6 Disability status scale	102
Appendix 7 ADL categories used within the Sonya Slifka dataset	103
References	104

Key definitions

Direct costs: The value of all goods, services and other resources that are consumed in the provision of an intervention or in dealing with the side effects or other current and future consequences linked to it.^(a)

Incremental cost-effectiveness ratio: The ratio of the difference in costs between two alternatives to the difference in *effectiveness* between the same two alternatives.^(a)

Indirect costs: Productivity gains or losses related to illness or death.^(a)

Marginal cost-effectiveness ratio: The incremental *cost-effectiveness ratio* between two alternatives that differ by one unit along some quantitative scale of intensity, dose, or duration.^(a)

Markov Chain Monte Carlo: A class of algorithms for sampling from probability distributions based on constructing a Markov chain.

Net benefit: The valuation of the benefits resulting from the use of a healthcare intervention in monetary terms, less any cost consequences. Where health gains are valued in terms of quality adjusted life years, this corresponds to the number of QALY produced by an intervention \times willingness to pay threshold (λ) – the cost of the providing the intervention.

Probabilistic sensitivity analysis: A form of model testing in which uncertainty surrounding all model parameter values are described by probability distributions rather than single (mean) values. This uncertainty is propagated through the model over a large number of iterations in order to generate information on the likelihood that a given intervention produces the greatest amount of net benefit.

Quality adjusted life year (QALY): A measure of health outcome which assigns to each period of time a weight, usually ranging from 0 to 1, corresponding to the *health-related quality of life* during that period, where a weight of 1 corresponds to optimal health, and a weight of 0 corresponds to a state of health judged equivalent to death.^(a) Survival gains are then weighted by these quality of life adjustments and aggregated to produce QALYs.

Willingness-to-pay: The maximum amount for which an individual is willing to pay to achieve a given level of health benefit (usually denoted by λ).

(a) Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. 1996. Oxford University Press: New York.

Abbreviations

ABN	Association of British Neurologists
ADL	Activities of Daily Living
C1MSSG	Copolymer-1 Multiple Sclerosis Study Group
CCTR	Cochrane Controlled Trials Register
CDSR	Cochrane Database of Systematic Reviews
CEAC	Cost-effectiveness Acceptability Curve
CMS	Centers for Medicare and Medicaid Services
CNS	Central Nervous System
COS	Community of Science
CSF	Cerebrospinal fluid
DSS	Disability Status Scale
EDSS	Expanded Disability Status Scale
EQ-5D	EuroQol 5D
EVIDENCE	Evidence of Interferon Dose-response: European North American Comparative Efficacy
FDA	Food and Drug Administration
GA	Glatiramer acetate
HRQoL	Health-Related Quality of Life
HUI3	Health Utilities Index (Mark 3)
IFNBMSSG	Interferon Beta Multiple Sclerosis Study Group
IFNβ	Interferon beta
IM	Intramuscular
INCOMIN	INdependent COMparison of INterferons
ITT	Intention to treat
MCMC	Markov Chain Monte Carlo
MMA	Medicare Modernization Act
MRC	Medical Research Council
MRDD	Medicare Replacement Drug Demonstration
MRI	Magnetic resonance imaging
MS	Multiple Sclerosis
MSCRG	Multiple Sclerosis Collaborative Research Group
NASPMS	North American Secondary Progressive Multiple Sclerosis
NHS EED	NHS Economic Evaluations Database
NICE	National Institute for Clinical Excellence
NRR	National Research Register
PPMS	Primary Progressive Multiple Sclerosis

PRISMS	Prevention of Relapses and Disability by Interferon β 1a Subcutaneously in Multiple Sclerosis
PRMS	Primary Relapsing Multiple Sclerosis
QALY	Quality Adjusted Life Year
RCT	Randomized Controlled Trial
RRMS	Relapsing Remitting Multiple Sclerosis
RPMS	Relapsing Progressive Multiple Sclerosis
SPMS	Secondary Progressive Multiple Sclerosis

Executive summary

Background

This report describes one of two studies undertaken to address a Congressional mandate to evaluate the cost-effectiveness of expanded drug coverage to Medicare under the Medicare Replacement Drug Demonstration. The demonstration aimed to improve beneficiary access to selected new oral anti-cancer drugs and other self-injected medications such as those used to treat multiple sclerosis (MS). Currently, treatment of MS using IFN β -1a is covered under Medicare Part B when administered intra-muscularly by a physician. Until the new prescription drug benefit (Medicare Part D) began in 2006, subcutaneous injection of IFN β -1a administered by the patient was not covered. The MRDD provided temporary national coverage to all self-administered treatments during the 16 months before Medicare Part D was implemented. Patient cost-sharing under the MRDD was structured to resemble Part D. The second study examines the cost-effectiveness of anti-TNF inhibitors for treating rheumatoid arthritis.

The main question addressed by this review is “*What is the clinical effectiveness and cost-effectiveness of interferon beta (IFN β) and glatiramer acetate (GA) in the management of relapsing/remitting multiple sclerosis and secondary-progressive multiple sclerosis to the Medicare program in the United States?*”

Clinical effectiveness methods

The relevant patient population for inclusion within the review of clinical effectiveness was:

- adults with RRMS, eligible for treatment with IFN β or GA; *or*
- adults with SPMS, eligible for treatment with IFN β .

Five interventions were included in the review:

- GA 20mg (Copaxone/Copolymer-1), daily subcutaneous injection;
- IFN β -1a 22 μ g (Rebif), subcutaneous injection 3 times a week;
- IFN β -1a 44 μ g (Rebif), subcutaneous injection 3 times a week;
- IFN β -1a 6MIU (Avonex), intramuscular injection once per week;
- IFN β -1b 8MIU (Betaseron), subcutaneous injection every other day.

The relevant comparator for the assessment was placebo, or another disease-modifying therapy where head-to-head studies were available. The following co-interventions were defined for inclusion within the review: best supportive care, including symptom control, physiotherapy, psychiatric and social support, disability aids, concomitant medication (not immunomodulatory therapy) for relapses or treatment-related adverse events.

The following clinical effectiveness outcomes were included within the review of clinical effectiveness:

- EDSS disease progression rates;
- Relapse rates;
- Health-related quality of life;
- Adverse events / toxicity;
- Study withdrawals and dropouts.

Studies for inclusion in the review included randomized controlled trials, which could be either placebo-controlled or head-to-head trials. Non-randomized trials would have been included in the review only if randomized studies were not available. Systematic reviews were also identified; while these were not included in the review of clinical effectiveness, these were retained for discussion. Studies were excluded from the review if off-label doses or off-label administrations of IFN β or GA were employed. Studies of other medications for MS not listed as included interventions were excluded. Clinical trials which did not report EDSS progression data were also excluded from the review.

Pre-specified outcomes were tabulated and discussed within a descriptive synthesis. Evidence on the comparative efficacy of three of the interventions included in the review (plus placebo) was synthesized using mixed treatment comparisons models.

Clinical effectiveness results

The systematic searches identified nine RCTs which met the inclusion criteria. Five of these studies were placebo-controlled trials in patients with RRMS, and two were placebo-controlled trials in patients with SPMS. The remaining two studies were head-to-head trials in patients with RRMS.

The randomization strategy was adequate in six of the placebo-controlled trials, and unclear in one trial. Randomization strategies were adequate in both head-to-head trials. Concealment of treatment assignment was adequate in five of the placebo-controlled trials. In all seven placebo-controlled trials, patients, providers and outcome assessors were blinded to treatment group. Blinding may have been compromised by adverse events in intervention groups (typically injection-site reactions). In all seven placebo-controlled trials, examining neurologists/physicians were separate from treating physicians. Patients and providers were not blinded in the head-to-head trials.

For the treatment of RRMS, all interventions significantly reduced relapse rates when compared with placebo. IFN β -1b 8MIU and IFN β -1a 44 μ g both reduced relapse rates significantly more

than IFN β -1a 6MIU. Based upon the analysis of the placebo-controlled RCTs, the relative reduction in relapse rate was estimated to be around 30% for IFN β -1a 44 μ g, IFN β -1a 22 μ g, IFN β -1b 8MIU and GA 20mg. The relative reduction in relapse rate was estimated to be approximately 18% for IFN β -1a 6MIU versus placebo. When evidence from the head-to-head trials was included in the analysis, mean reductions in relapse rates appear very similar to the placebo-controlled trial estimates. IFN β -1b 8MIU significantly improved relapse rates compared with placebo in the treatment of SPMS; the relative reduction in relapse rate for IFN β -1b 8MIU versus placebo was estimated to be between 26% and 43%.

For placebo-controlled trials in patients with RRMS, significant improvements in time to sustained disease progression were reported for IFN β -1a at both 44 μ g and 22 μ g doses, and for IFN β -1a 6MIU compared against placebo. The latter study was terminated prematurely, and many reported analyses were not undertaken according to the intention-to-treat principle. The relative hazard ratio for progression for GA 20mg versus placebo was estimated to be in the range 0.76 to 0.86. The estimated relative hazard ratio for progression for IFN β -1a 6MIU versus placebo was estimated to be 0.58. The relative hazard ratio for progression for IFN β -1b 8MIU versus placebo was estimated to be 0.71. The estimated relative hazard ratios for progression for IFN β -1a 44 μ g and IFN β -1a 22 μ g versus placebo were estimated to be 0.60 and 0.72 respectively. Estimates of uncertainty surrounding these hazard ratios were not available (or calculable) from the limited data reported within the study publications.

A significant improvement in time to sustained EDSS progression was reported for IFN β -1b 8MIU compared to IFN β -1a 6MIU. IFN β -1a 44 μ g did not significantly delay time to disease progression compared to IFN β -1a 6MIU, although the data suggested a progression risk reduction in favor of the IFN β -1a 44 μ g group. The mixed treatment comparisons model suggested that the modified relative hazard ratios for disease progression for IFN β -1a 6MIU, IFN β -1b 8MIU, and IFN β -1a 44 μ g versus placebo were 0.79, 0.52, and 0.70 respectively.

For the treatment of SPMS, one study reported a statistically significant improvement in time to disease progression for IFN β -1b 8MIU; a non-significant difference was reported for the other SPMS trial. The relative hazard ratio for progression for IFN β -1b 8MIU versus placebo was estimated to be in the range 0.72 to 0.93.

Health economic methods

Evidence on the costs and effects of the use of IFN β and GA in the treatment of MS were synthesised within a state transition cohort model. The model was originally developed for use in the appraisal of IFN β and GA on behalf of the National Institute for Clinical Excellence in England and Wales in 2001. As with the earlier UK cost-effectiveness analysis, the assumptions

employed within the model developed for CMS favor IFN β and GA over best supportive care. The original cost-effectiveness model was updated to incorporate a lifetime horizon for the evaluation of the costs and effects of alternative treatment options, the incorporation of US-specific data on baseline disease severity, updated estimates of treatment effectiveness derived from the review of clinical effectiveness, updated estimates of health-related quality of life by EDSS state from a cohort of patients within Nova Scotia, and US-specific data on the costs of disease-modifying therapies and the costs of MS care from the perspective of the health care payer. The population for the cost-effectiveness analysis is the representative Medicare beneficiary with MS, as described by and agreed with the CMS.

Health economic results

Under the base case assumptions the marginal cost-effectiveness of the disease-modifying therapies compared to best supportive care is expected to be greater than \$100,000 per QALY gained. The central estimate of cost-effectiveness for IFN β -1a 6MIU versus best supportive care is estimated to be approximately \$112,500 per QALY gained. When this therapy is self-administered, this results in a slightly more favorable estimate of \$104,200 per QALY gained. However, the trial from which the efficacy estimate for IFN β -1a 6MIU versus placebo was derived was terminated early and the clinical analysis was not undertaken according to the ITT principle. The central estimates of cost-effectiveness for IFN β -1a 22 μ g and IFN β -1a 44 μ g versus best supportive care are estimated to be \$198,500 per QALY gained and \$131,900 per QALY gained respectively. For the treatment of RRMS, the central estimate of cost-effectiveness for IFN β -1b 8MIU versus best supportive care is estimated to be \$164,100 per QALY gained. When disease-modifying therapy is assumed to continue upon progression to SPMS, the central estimate of cost-effectiveness for IFN β -1b 8MIU versus best supportive care is estimated to be \$295,200 per QALY gained. The central estimate of cost-effectiveness for GA 20mg versus best supportive care is estimated to be \$332,000 per QALY gained.

When information from the head-to-head trials is used to modify the estimated effectiveness of the disease-modifying therapies, the marginal cost-effectiveness of IFN β -1a 6MIU is estimated to be in the range \$230,800 - \$248,600 per QALY gained when compared against best supportive care. The results of the head-to-head analysis of IFN β -1a 6MIU are broadly consistent with the UK commercial-in-confidence ITT analysis, whereby the marginal cost-effectiveness ratio for IFN β -1a 6MIU versus best supportive care appears considerably less economically attractive than the cost-effectiveness estimate produced using the public domain placebo-controlled trial estimates of effectiveness. When the model includes information from the head-to-head trials, the adjusted marginal cost-effectiveness of IFN β -1b 8MIU is estimated to be approximately \$91,100 per QALY gained compared to best supportive care. Based on the modified treatment effectiveness estimates, the marginal cost-effectiveness of IFNB-1a 44 μ g versus best supportive care is

estimated to be \$179,900 per QALY gained. However, owing to heterogeneities between the trials in terms of patients and the definitions of progression and relapse endpoints, the results of these analyses should be interpreted cautiously.

Owing to the absence of US clinical guidelines concerning treatment cessation, the base case model assumes that all patients continue treatment until EDSS 10. However, if one assumes that patients discontinue therapy upon progression to EDSS 7.0, the analysis produces notably more favorable marginal cost-effectiveness estimates for all of the disease-modifying therapies compared to best supportive care. This finding is of particular relevance to the MS Medicare population in the US, who on average are likely to be older and more disabled than those patients recruited into the clinical trials from which estimates of efficacy were drawn. The sensitivity analysis shows that the central estimates of cost-effectiveness presented here are subject to considerable uncertainty, particularly with respect to the long-term impact of these therapies on clinical outcomes. The simple sensitivity analysis also demonstrates that the effectiveness of disease-modifying therapy upon EDSS progression has a considerable impact upon the cost-effectiveness of these therapies, while the impact of therapy upon relapse does not substantially affect cost-effectiveness outcomes. This does not mean that relapses are not clinically important, but rather that the impact of treatment upon disease progression has greater capacity to affect the cost-effectiveness of these therapies.

Under the base case model assumptions whereby patients are assumed to continue treatment until EDSS 10, the probability that any of the disease-modifying therapies has a marginal cost-effectiveness that is better than \$60,000 per QALY gained when compared to best supportive care is estimated to be 0.10 or lower. The probability that any of the disease-modifying therapies has a marginal cost-effectiveness that is better than \$100,000 per QALY gained when compared to best supportive care is estimated to be 0.48 lower. On account of the substantial improvement in cost-effectiveness demonstrated when all patients were assumed to stop treatment upon progression to EDSS 7.0, further probabilistic sensitivity analysis was undertaken for this scenario. Under this assumption, the probability that any of the disease-modifying therapies has a marginal cost-effectiveness that is better than \$60,000 per QALY gained when compared to best supportive care is estimated to be 0.50 or lower. The probability that any of the disease-modifying therapies has a marginal cost-effectiveness that is better than \$100,000 per QALY gained when compared to best supportive care is estimated to be 0.76 or lower.

Conclusions

This assessment is intended to represent the most appropriate and robust analysis of the cost-effectiveness of the disease-modifying therapies for the US Medicare MS population given current evidence. However, the current state of evidence concerning the chronic use of these therapies is

subject to several important gaps and uncertainties. Where evidence is weak or absent, the assumptions employed within the cost-effectiveness analysis favor the disease-modifying therapies. Under the base case model assumptions, the cost-effectiveness of IFN β -1a, IFN β -1b and GA is in excess of \$100,000 per QALY gained when compared against best supportive. Cessation of treatment upon progression to EDSS 7.0 may provide a more cost-effective treatment strategy for all of the therapies.

Areas for further research

The state of knowledge on the clinical and cost implications of the disease-modifying therapies is continuously evolving. The review of clinical effectiveness and the development of the US cost-effectiveness model for the disease-modifying therapies highlights a number of areas in which further research is warranted.

- Existing RCTs of IFN β and GA have used trial durations of between 9 months and 5 years. Further research concerning the impact of the disease-modifying therapies on long-term disease progression and relapse is required.
- There is a dearth of evidence concerning the effectiveness of sequences of disease-modifying therapies for MS. Further research concerning the clinical effectiveness and cost-effectiveness of alternative sequences of disease-modifying therapies is merited.
- Further research on the relationship between the EDSS, health care utilization and costs of care in the US Medicare MS population would be valuable.
- Finally, research concerning the relationship between the EDSS and health utility in the US population may enhance the external validity of subsequent cost-effectiveness analyses of MS therapies.

1.0 Aims and objectives

1.1 Background to study

This report describes one of two studies undertaken to address a Congressional mandate to evaluate the cost-effectiveness of expanded drug coverage to Medicare under the Medicare Replacement Drug Demonstration (MRDD). The demonstration aimed to improve beneficiary access to selected new oral anti-cancer drugs and other self-injected medications such as those used to treat multiple sclerosis (MS). Currently, treatment of MS using IFN β -1a is covered under Medicare Part B when administered intra-muscularly by a physician. Until the new prescription drug benefit (Medicare Part D) began in 2006, subcutaneous injection of IFN β -1a administered by the patient was not covered. The MRDD provided temporary national coverage to all self-administered treatments during the 16 months before Medicare Part D was implemented. Patient cost-sharing under the MRDD was structured to resemble Part D. The second study examines the cost-effectiveness of anti-TNF inhibitors for treating rheumatoid arthritis.

1.2 Study aim

The main question addressed by this review is “*What is the clinical effectiveness and cost-effectiveness of interferon beta (IFN β) and glatiramer acetate (GA) in the management of relapsing/remitting multiple sclerosis (RRMS) and secondary-progressive multiple sclerosis (SPMS) to the Medicare program in the United States?*”

1.3 Study objectives

More specifically the objectives of the study are:

- To evaluate the clinical effectiveness of IFN β and GA in the management of RRMS and SPMS within their label indications as compared with best supportive care or Avonex;
- To estimate the cost-effectiveness of IFN β and GA in the management of RRMS and SPMS as compared with best supportive care or Avonex from the perspective of the US health care payer.

1.4 Structure of the report

Chapter 2 presents a background to the natural history and diagnosis of MS, and describes the alternative pharmacological and non-pharmacological approaches to the management of the disease.

Chapter 3 reports the methods for the systematic review of clinical effectiveness of IFN β and GA in the management of MS. In addition, methods for estimating effectiveness parameters for use in the cost-effectiveness model are described in detail.

Chapter 4 presents the results of the systematic review of clinical effectiveness. Effectiveness parameters used in the subsequent cost-effectiveness analysis (i.e. relative relapse rates and relative hazard ratios for MS progression) are also reported within this chapter of the report.

Chapter 5 presents the methods used to evaluate the cost-effectiveness of IFN β and GA in the management of MS. All major structural assumptions and sources of evidence used within the model are described within this chapter.

Chapter 6 reports the results of the cost-effectiveness analysis. The base case analysis reports estimates of the marginal cost-effectiveness of each therapy compared to best supportive care based exclusively upon the placebo-controlled evidence. Simple sensitivity analyses are presented to demonstrate the impact of discontinuing treatment upon progression to EDSS 7.0, alternative model time horizons, assumptions concerning the relationship between MS disability and the costs of care, and alternative values for other model parameters on the central estimates of cost-effectiveness. A further sensitivity analysis is presented using modified treatment effectiveness estimates informed by the head-to-head trials of disease-modifying therapies. All health economic analyses are reported from the perspective of the US health care payer. Finally, this chapter reports the results of the probabilistic sensitivity analysis, as represented by marginal Cost Effectiveness Acceptability Curves (See “Key Definitions”).

Chapter 7 discusses the results of the assessment of clinical effectiveness and cost-effectiveness and highlights implications for clinical practice in the US. Limitations concerning the current state of clinical and cost evidence are highlighted and their implications on the cost-effectiveness of the disease-modifying therapies are discussed. Areas in which further research is indicated are also discussed.

Chapter 8 presents the conclusions on the clinical effectiveness and cost-effectiveness of the disease-modifying therapies in the management of MS.

2.0 Background

2.1 Prevalence of MS

MS is a demyelinating disease of the central nervous system (CNS). MS causes damage to the myelin nerve fibre in the brain and spinal cord through the development of hardened sclerotic plaques, which in turn interrupt the transmission of nerve impulses. MS is one of the most common neurological conditions affecting young adults, and is two to three times more common in women than men.^{1,2} An estimated 400,000 people in the US have MS, and its prevalence in the Northern states is reported to be higher than in the Southern states.²

2.2 Clinical course of MS

Three common classifications of MS have been identified according to the characteristics of the underlying disease: relapsing/remitting MS (RRMS); secondary progressive MS (SPMS); and primary progressive MS (PPMS). The broad clinical course of these types of MS is detailed in Table 1.

Table 1 Classifications of MS with general descriptions of disease characteristics³

MS classification	Description of disease characteristics
<i>Common disease classifications</i>	
Relapsing/remitting MS	Clearly defined disease relapses with full recovery or with sequelae and residual deficit upon recovery; periods between relapses characterized by a lack of disease progression. Approximately 80% of patients have RRMS at onset.
Primary progressive MS	Disease progression from onset with occasional plateaus and temporary minor improvements. Approximately 10-15% patients have PPMS disease at onset.
Secondary progressive MS	Initial relapsing-remitting course followed by progression with or without occasional relapses, minor remissions and plateaus. Approximately 50% of people with RRMS develop SPMS during the first 15-20 years of their illness.
<i>Less common disease classifications</i>	
Progressive relapsing MS (PRMS)	Progressive from onset with clear acute relapses but with progression between relapses (this is probably comparable to primary progressive disease).
Benign MS	Disease in which the patient remains fully functional in all neurologic systems.
Malignant MS	Rapid progressive course leading to significant disability in multiple neurologic systems or death in a relatively short time following disease onset.

These disease classifications remain controversial given the limited epidemiological evidence against which to assess their validity or reliability. Attempts to refine these MS classifications have led to difficulty in combining estimates of prevalence and incidence across studies.⁴

2.3 Diagnosis of MS

MS is usually diagnosed between the ages of 20 and 50 years.¹ While the etiology of MS is currently unclear, there is evidence to suggest that MS results from an autoimmune response, although it has been suggested that both environmental and genetic factors play a role.¹ MS is diagnosed clinically, and is dependent upon obtaining evidence of sclerotic lesions located in different parts of the CNS at different points in time.¹ There is no definitive diagnostic test for MS; a confirmed diagnosis of established MS may require several neurological tests and clinical assessment over a period of time. It may not be possible to make a definite diagnosis of MS during the early stages of the disease. MS is typically diagnosed through clinical assessment, magnetic resonance imaging (MRI) or through examination of cerebrospinal fluid (CSF). However even among specialist neurologists there may be disagreement.³ In 2001, the Royal College of Physicians (London, UK) issued the "McDonald Criteria" for the diagnosis of MS;⁵ these criteria integrate MRI assessment with clinical and other paraclinical methods. The McDonald Criteria have been extensively used and assessed; revisions to the criteria have subsequently been published.⁶

2.4 Symptoms associated with MS

MS is characterized by a variety of debilitating symptoms including pain, fatigue, muscle spasticity and spasm, ataxia and tremor, bladder disturbance and micturition problems, sleep disturbance, balance and postural problems, optic neuritis, paraesthesiae, diplopia and vertigo. In general, the severity of MS symptoms reflect the degree of myelin loss that has taken place, although the development of sclerotic patches leads to variation in symptoms experienced between patients, as well as variation in the symptoms experienced by an individual patient throughout the course of the disease.¹

2.5 Measurement of disease progression

Disease progression is typically measured in terms of impairment and disability, using Kurtzke's Expanded Disability Status Scale (EDSS), an ordinal scale ranging from EDSS 0 (normal neurologic examination) to EDSS 10 (death due to MS).⁷ The EDSS is presented in Appendix 1. Up to EDSS 3.5 the scale measures neurological impairments that are likely to have limited if any impact upon the activities of daily living. EDSS scores between 4.0 and 5.5 reflect ambulatory limitations. At EDSS 6.0 and EDSS 6.5, patients require intermittent or constant mobility assistance. For scores over EDSS 7.0, patients will require a wheelchair. The progression to SPMS normally takes place over the EDSS range 2.5 to 4.5. Disability progression is associated with permanent reductions in quality of life and increases in the cost of medical management.⁷

2.6 Burden of disease

The burden of MS is substantial, as the disease affects every aspect of a sufferer's life including self care, relationships, work and travel. The burden for families and carers is also considerable. There is no cure for MS, hence a central focus of the management of MS concerns improving the patient's day-to-day quality of life by treating those symptoms which have the greatest impact on the physical, social and emotional wellbeing of patients, their families and their carers. The emotional, social and physical burden of MS is well documented in anecdotal patient reports and surveys.

As the disease progresses and the patient's symptoms become more severe, greater healthcare resources are required for the management of the disease. The annual per patient cost varies considerably according to the progression of the disease and the severity of symptoms experienced. The financial burden of MS is considerable, although published evidence on the cost of managing the disease in the US is limited. Whetten-Goldstein and colleagues⁸ undertook an analysis of costs of MS in the US based upon survey data and various secondary sources. The analysis included the costs of personal health services such as the use of hospital resources, physician costs, drug costs, personal assistance, formal and informal care, domestic help, occupational therapy, and lost earnings. The authors estimated the annual cost of MS to be \$34,000 per person, which translated to a lifetime per patient cost of \$2.2 million (1994 US prices). The analysis included both direct and indirect costs; lost earnings and the costs of informal care were identified as major cost components. The annual cost of MS within the US was estimated to be \$6.8 billion.⁸

2.7 Current treatments available for the management of MS

Disease management focuses on slowing progression and preventing relapse as well as symptom-control. It has been suggested that patients are principally concerned with limiting their handicap by attention to social, vocational, marital, sexual and psychological aspects of the disease.¹ A recent systematic review of pharmacological and non-pharmacological therapies used in the management of MS spasticity and pain⁹ identified a large number of studies concerning 15 pharmacological therapies used to manage MS spasticity and 14 pharmacological therapies used to manage MS pain.

The variation in symptoms experienced by patients throughout the course of disease, together with the vast range of pharmacological interventions used to manage these symptoms (See Table 2) highlights that there is currently no effective standard current therapy for the universal management of MS. Further, the efficacy of many therapies may be limited by their toxicity.⁹ Evidence concerning the effectiveness of many of these therapies is very limited, thus many interventions are commonly prescribed off-label.⁹ Non-pharmacological treatments are also used

to manage MS symptoms; these may include speech therapy, physiotherapy, and occupational therapy.

Table 2 Pharmacological therapies in the management of symptoms association with MS¹

Symptom	Therapies used to treat early disease	Therapies used to treat late disease
Spasticity	Baclofen, dantrium, i.v. methylprednisolone, benzodiazepines, threonine, tizanadine, vigabatrin, chlonodine, mexilitinem ivermectin, cannabinoids	Intrathecal baclofen, botulinum toxin, phenol, tendon surgery, rhizotomy, nerve section, magnetic stimulation
Tremor	β blockers, primidone, glutethamide, clonazepam, izoniazid, ondanestron, hyocsine, carbamazepine, sodium valproate	Streotactic thalamotomy, thalamic stimulation, physical restraint
Dizziness and nystagmus	Cinnarizine, prochorperazine, memantine	
Fatigue	Amantidine, permoline, fluoxetine	
Strength		4-aminopyridine, 2,4-diaminopyridine, electrical stimulation
Bladder storage (detrusor)	Oxybutynin, propantheline, imipramine, flavoxate, dicyclamide, maprotilene, isoprenaline, empromium	
Bladder storage (sphincter)	Ephedrine, phenylpranolamine, imipramine, desmopressin	
Bladder emptying (detrusor)	Carbachol, bethanecol, distigmine, abdominal pressure	Local phenol, capsaicin, lidocaine, verapamil
Bladder emptying (sphincter)	Phenoxybenzamine, prazcin, terazocin, diazepam, baclofen, dantrium, perineal stimulation	Electrical stimulation, bladder neck surgery
Combined storage and emptying	Self catheterization \pm oxybutinin	Artificial sphincters, permanent catheterization, urinary diversion
Bowel	Loperamide, bulk laxatives, enemas	Faecal containment, colostomy
Sexual (males)	Papaverine, yohimbine, phentolamine, sildenafil, prostaglandin E ₁	Mechanical prostheses, electro-ejaculation
Sexual (females)	Artificial lubrication	
Paroxysmal	Carbamazepine, other anti-convulsants, misoprostol	
Pain	Anti-convulsants, anti-depressants, non-steroidal drugs	Nerve section, alcohol injection, sympathetic block, cutaneous nerve stimulation, TENS

2.8 Disease-modifying therapies for the management of MS

The class of drugs known as “disease-modifying therapies”, namely the interferon betas (Rebif[®] - Serono, Betaseron[®] - Schering Healthcare Ltd., Avonex[®] - Biogen) and glatiramer acetate

(Copaxone[®] - TEVA/Aventis) entered the market during the 1990s. These therapies are thought to slow the progression of the disease, and reduce the number and severity of relapses experienced. The indications of these therapies which have been approved by the US Food and Drugs Administration (FDA) are shown in Table 3. More recently, Mitoxantrone (Novantrone[®] - Serono), an additional disease-modifying therapy has also become available for the management of MS. However, reimbursement for Novantrone was not covered under the MRDD by the CMS and is thus not considered within this review.

Table 3 Disease-modifying therapies for the management of MS

Product name	Drug	Manufacturer	Dosage	Label indications
Avonex	IFN β -1a	Biogen	6MIU once a week	RRMS
Betaseron	IFN β -1b	Schering	8MIU every other day	RRMS and SPMS
Rebif	IFN β -1a	Serono	22 μ g or 44 μ g three times a week	RRMS
Copaxone	GA	TEVA/Aventis	20mg daily	RRMS
Novantrone	Mitoxantrone	Serono	Usually 12mg/m ² once every 3 months (8-12 doses)	RRMS, SPMS and PRMS

Details of indications of these disease-modifying therapies are given below.

2.8.1 Avonex (IFN β -1a, Biogen)¹⁰

Avonex has been approved by the FDA for the treatment of patients with relapsing forms of MS to slow the accumulation of physical disability and decrease the frequency of clinical exacerbations. Patients with MS in whom efficacy has been demonstrated include those who have experienced a first clinical episode and have MRI features consistent with MS. Avonex is given as a 6MIU once-a-week intramuscular (IM) injection, usually in the large muscles of the thigh, upper arm, or hip. Avonex can also be administered subcutaneously, although this indication has not been approved by the FDA.

2.8.2 Betaseron (IFN β -1b, Schering)¹⁰

Betaseron has been approved by the FDA for the treatment of relapsing forms of MS to reduce the frequency of clinical exacerbations. Relapsing forms of MS include individuals with SPMS who continue to experience relapses or acute attacks. Betaseron is injected subcutaneously (between the fat layer just under the skin and the muscles beneath) every other day at a dose of 8MIU.

2.8.3 Rebif (*IFN β -1a, Serono*)¹⁰

Rebif has been approved for the treatment of patients with relapsing forms of MS to decrease the frequency of clinical exacerbations and delay the accumulation of physical disability. Low-dose (22 μ g) or high-dose (44 μ g) Rebif is given three times a week subcutaneously (between the fat layer just under the skin and the muscles beneath).

2.8.4 Copaxone (*glatiramer acetate, TEVA*)¹⁰

GA has been approved by the FDA to reduce the frequency of relapses in patients with RRMS. A dose of 20mg GA is injected subcutaneously (between the fat layer just under the skin and the muscles beneath) once a day.

2.8.5 Novantrone (*Mitoxantrone, Serono*)¹⁰

Novantrone has been approved by the FDA for reducing neurologic disability and/or the frequency of clinical relapses in:

1. Patients with SPMS (disease that has changed from RRMS to progressive MS at a variable rate);
2. Patients with PRMS (disease characterized by gradual increase in disability from onset with clear, with acute relapses along the way);
3. Worsening RRMS (disease characterized by clinical attacks without complete remission, resulting in a step-wise worsening of disability).

2.9 Provision of disease-modifying therapies under Medicare

Prior to 2002, Medicare coverage for IFN β s to treat MS was sporadic, varying by carrier depending on their interpretation of whether a drug was usually self- or physician administered. The manufacturers of the disease-modifying therapies have implemented programmes designed to help individuals apply for these therapies, in some cases, where patients are uninsured or under-insured. In 2002, CMS clarified the definition of self-administered drugs, allowing for coverage of IFN β -1a for treatment of MS if administered intra-muscularly by a physician. Subcutaneous injection of IFN β -1a administered by the patient was not covered. The MRDD expanded coverage to all self-administered treatments nationwide.

3.0 Methods for the review of clinical effectiveness

3.1 Search strategy

Systematic searches were undertaken to identify all studies relating to IFN β and GA in the management of RRMS and SPMS. Search terms included generic and commercial drug names (for example, interferon beta, Betaseron/Betaferon, beta interferon, Avonex, glatiramer acetate, Copaxone, copolymer-1, Rebif, interferon beta-1a, interferon beta-1b). The systematic searches were undertaken using the following electronic databases: MEDLINE, EMBASE, the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Controlled Trials Register (CCTR), the Science Citation Index and the NHS Centre for Reviews and Dissemination databases (DARE, NHS EED, HTA). PRE-MEDLINE was also searched in order to identify any studies not yet indexed on Medline. Current research was searched via The National Research Register (NRR), the Community of Science (COS) Funded Research database, the Current Controlled Trials Register and the MRC Clinical Trials Register. Any relevant systematic reviews were hand-searched in order to identify any further clinical trials. Searches were not restricted by language, date or publication type.

3.2 Inclusion criteria

3.2.1 Patients

The relevant patient population for inclusion within the review of clinical effectiveness was:

- adults with RRMS, eligible for treatment with IFN β or GA; *or*
- adults with SPMS, eligible for treatment with IFN β .

3.2.2 Interventions

Five interventions were included in the review:

- GA 20mg (Copaxone/Copolymer-1), daily subcutaneous injection;
- IFN β -1a 22 μ g (Rebif), subcutaneous injection 3 times a week;
- IFN β -1a 44 μ g (Rebif), subcutaneous injection 3 times a week;
- IFN β -1a 6MIU (Avonex), intramuscular injection once per week;
- IFN β -1b 8MIU (Betaseron), subcutaneous injection every other day.

3.2.3 Comparators

The relevant comparator for the assessment was placebo, or another disease-modifying therapy in instances where head-to-head trials were available. The following co-interventions were defined for inclusion within the review: best supportive care, including symptom control, physiotherapy,

psychiatric and social support, disability aids, concomitant medications (specifically excluding other immunomodulatory therapies) for relapses or treatment-related adverse events.

3.2.4 Outcome measures

The following outcome measures were included in the review of clinical effectiveness:

- EDSS disease progression rates;
- Relapse rates;
- Health-related quality of life (HRQoL);
- Adverse events/treatment-related toxicities;
- Study withdrawals and dropouts.

3.2.5 Methodology

There remains controversy concerning the role of observational studies to produce unbiased estimates of treatment effectiveness. Traditionally, randomized controlled trials (RCT) are considered to represent the “gold standard” in study design, as the experimental method of randomly assigning study participants between groups results in comparable groups which are balanced in terms of known, unknown and unmeasured confounding variables.¹¹ Observational studies, for the most part, are considered to be under greater threat of internal bias and confounding. While some empirical evidence in other disease areas suggests that observational data may produce highly similar estimates of effectiveness to those obtained from experimental RCT designs, other comparative studies have demonstrated that observational data may be distorted in either direction. Consequently, it may be impossible to predict whether the results of observational studies have been biased by uncontrolled or unbalanced factors.¹¹

It is noteworthy that long-term observational studies are available for some of the disease-modifying therapies used to manage MS. For example, 10-year outcomes for MS patients receiving glatiramer acetate have recently been published.¹² However, such observational data are non-randomized, uncontrolled, and are likely to be subject to a considerably greater degree of bias than properly designed RCTs. Therefore, this review included only RCTs, which could be either placebo-controlled or head-to-head trials. Non-randomized studies would have been included in the review *if* randomized studies were not available. Systematic reviews were also identified from the searches; while these studies were not included in the review of clinical effectiveness, these were retained for discussion.

3.3 Exclusion criteria

Studies were excluded if off-label doses or administrations of IFN β or GA were employed. Studies of other medications for MS not listed as included interventions were excluded (See

Section 3.2.2). Clinical trials which did not report EDSS progression data were also excluded from the review.

3.4 Data extraction strategy

Data were extracted by one researcher, and checked by a second, using a standardized data extraction form. Any disagreements were resolved by discussion.

3.5 Quality assessment strategy

As noted in Section 3.3, only RCTs were included within the review of clinical effectiveness. On account of the availability of RCT evidence, non-randomized studies were excluded from the review. The methodological quality of RCTs were assessed through the evaluation of four recognized components of trial quality; generation of randomization schedule, concealment of treatment assignment, blinding and exclusion of patient data from analyses.¹³ A narrative account of trial quality was produced.

3.6 Clinical review methods

3.6.1 Analysis and synthesis of clinical effectiveness evidence

Pre-specified outcomes were tabulated and discussed within a descriptive synthesis. Evidence on the comparative efficacy of three of the interventions included in the review (plus placebo) was synthesized using mixed treatment comparisons models.

3.6.2 Methods for estimating relative hazards of disease progression

The health economic model described in Chapter 5 simulates the natural history of a cohort of individuals with RRMS and SPMS. The effect of treatment on natural history disease progression was modeled as a relative hazard ratio for patients receiving each disease-modifying therapy compared to placebo (or for head-to-head trials, compared to some other disease-modifying therapy). Under best supportive care, the relative hazard ratio will be 1.0, whereas a lower hazard ratio was estimated for each of the treatments to represent the observed effects of slowing disease progression whilst on therapy. The magnitude of this relative hazard ratio for disease progression differs between each of the disease-modifying therapies. For the most part, relative hazard ratios for the disease-modifying therapies compared to placebo were not directly reported within the trial publications. In such instances, relative hazard ratios for disease progression were estimated using Kaplan Meier progression-free survival curves or using relevant narrative data reported in the text of the trial publications.

The methods for estimating relative hazard ratios for disease progression between treatment groups were based upon the assumption that EDSS progression-free survival follows an

exponential distribution. The transformation of the exponential survivor function $S(t)$ to derive relative hazard rates is shown below. This approach is analogous to using exponential regression analysis to estimate the relative hazard ratio based on Kaplan-Meier progression-free survival curves reported within the clinical trials.

Assumed distribution of progression free survival for disease-modifying therapy:

$$S(t) = e^{-a_1 t}$$

$$\Rightarrow \ln S(t) = -a_1 t$$

$$\Rightarrow a_1 = -\frac{1}{t} \ln \frac{1}{S(t)}$$

Assumed distribution of progression free survival for placebo:

$$\Rightarrow \ln S(t) = -a_2 t$$

$$\Rightarrow a_2 = -\frac{1}{t} \ln \frac{1}{S(t)}$$

Where $S(t)$ is the progression-free survivor function over time t , and a is the relative hazard rate of disease progression. Thus the relative hazard ratio for disease progression for disease-modifying therapy compared to placebo is calculated as: $\frac{a_1}{a_2}$

For illustration, a worked example of these calculations is presented in Appendix 2. It should be noted that the health economic model assumes that the relative hazard ratios for disease progression estimated from the placebo-controlled trials are maintained for the entire model time horizon (the remaining lifetime of the patient). In other words, the relative hazard ratio is assumed to neither increase nor deteriorate over time. However, the study durations employed within the randomized phase of the clinical trials of IFN β and GA ranged from 9 months¹⁴ to 5 years.¹⁵ The true impact of these therapies on disease progression and relapse over the course of the disease is subject to considerable uncertainty. This assumption favors *all* of the disease-modifying therapies over best supportive care, yet as the model time horizon increases, the validity of the exponential extrapolation becomes subject to an increasing degree of uncertainty.

3.6.3 Methods for estimating relative relapse rates

The health economic model simulates the annual number of relapses experienced within each EDSS state for patients receiving best supportive care. The relative relapse rates for each of the disease-modifying therapies compared to placebo were calculated using annualized relapse rates for each treatment group reported within the clinical trials included in the clinical effectiveness review. It should be noted that the annual relapse rate is *not* the same as a relative risk of relapse. Relative risks relate to a binary outcome (either [a] study subject experiences an event one or more

times *or* [b] study subject does not experience the event at all), while a relative rate incorporates information concerning the *number* of events experienced. This distinction is important, as patients may experience more than one relapse within a year. Annual relative relapse rates between treatment groups were estimated using the following formula.

$$\text{Relative relapse rate} = \frac{\text{annual number of relapses on disease-modifying therapy}}{\text{annual number of relapses on placebo}}$$

3.6.4 Methods for synthesizing evidence on hazard rates for disease progression

An exponential time-to-event model was fitted to progression data reported within the trials included in the clinical effectiveness review. With this model, the probability p that any given patient will progress within a follow-up period t is given by the following formula:

$$p = 1 - e^{-at}$$

where $1/a$ is the mean time to progression.

Each of the disease-modifying therapies has a different impact upon α . None of the trials have included arms for each treatment to enable a direct assessment of this impact. Instead, available progression rate data were synthesized using a mixed treatment model using WinBUGS software (Medical Research Council and Imperial College, London, 2003). This software allows for the practical application of Markov Chain Monte Carlo (MCMC, See “Key Definitions”) sampling methods. The WinBUGS progression model included random effects term which allows for between-trial heterogeneities. An explanation of the WinBUGS progression model is given below.

Let $d[j]$ be the impact of treatment j on $\log(\alpha)$ relative to placebo. Also, let $m[i]$ be the value of $\log(\alpha)$ for patients from study i if given placebo. Then the value of α (the hazard rate for an individual treatment arm) for study i , arm k , is given by:

$$\log(\alpha_{ik}) = m[i] + d[j_{ik}]$$

where j_{ik} is the index for the treatment given in arm k of study i .

The WinBUGS progression model syntax and data inputs are contained in Appendix 3.

3.6.5 Methods for synthesizing evidence on relapse rates

Evidence on relapse rates observed within the placebo-controlled trials and head-to-head trials were also synthesized within a separate mixed treatment comparison using WinBUGS software.

The mixed treatment comparison assumes that the number of relapses for patients within one arm of each trial is independently Poisson distributed. Under this assumption, it follows that the observed mean number of relapses in a trial arm estimates the expected mean number of relapses, and the observed mean number of relapses divided by the sample size in an arm estimates the variance of the mean number of relapses. Suppose we wish to estimate the relative relapse rate for IFN β -1a 6MIU versus placebo from a trial. The log mean relapse rate is a more suitable scale to work on, as this will be approximately normally distributed. The observed log mean relapse rate in each arm will be normally distributed with means q_a and q_p and variances estimated by the inverse of the product of the sample sizes and the mean relapse rates themselves, call these w_a and w_p . As the two trial arms are independent, the difference in log mean relapse rate (the log of the relative relapse rate), will be normally distributed with the mean $q_{ap} = q_a - q_p$ and variance estimated by $w_a + w_p$.

A similar exercise was repeated for each of the trials included in the review, and the five observed comparisons are given five normal distributions, the mean of each being the population log relative relapse rate for the comparison made within the trial. We are interested in estimating q_{ap} , q_{bp} and q_{rp} . Two of the comparisons (IFN β -1b 8MIU versus IFN β -1a 6MIU, and IFN β -1a 44 μ g versus IFN β -1a 6MIU) give estimates of q_{ba} and q_{ra} . These can be used to provide estimates of the comparisons we are interested in, as $q_{ba} = q_{bp} - q_{ap}$ and $q_{ra} = q_{rp} - q_{ap}$. MCMC methods¹⁶ were used to obtain estimates of the parameters $\exp(q_{ap})$, $\exp(q_{bp})$ and $\exp(q_{rp})$, the population relative relapse rates. The methods used to perform this data synthesis are described in further detail by Lu and Ades.¹⁷ The WinBUGS relapse model was a fixed effect indirect comparison meta-analysis and did not allow for between-trial heterogeneities due to the limited data reported in the trial publications.

The WinBUGS relapse model syntax and data are detailed in Appendix 3.

4.0 Clinical effectiveness results

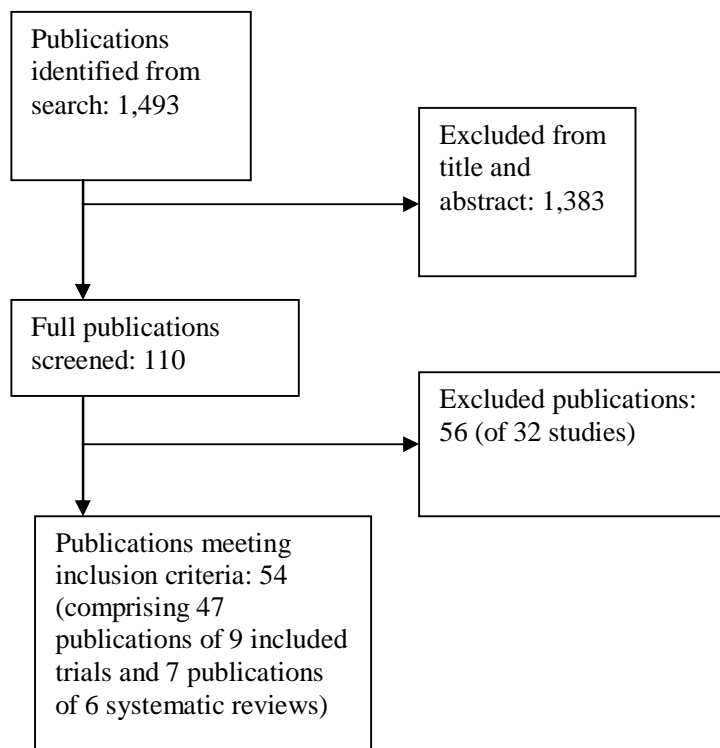
4.1 Number of studies identified

The electronic searches identified a total of 1,493 publications. The titles and abstracts of all publications were reviewed, which led to the exclusion of 1,383 publications. Publications were excluded from the review for the following reasons:

- study designs which did not meet the inclusion criteria set out in Section 3.2.5 (n=948);
- duplicate references (n=265);
- non-English language publications (n=154);
- irrelevant intervention (n=8);
- irrelevant patient group (n=8).

One hundred and ten publications were retrieved for further detailed evaluation. These comprised 47 publications of 9 included trials; 56 publications of 32 excluded studies; and 7 publications of systematic reviews. A summary of the publications included and excluded from the review is detailed in the Quality of Reporting of Meta-Analyses (QUOROM)¹⁸ flowchart shown in Figure 1. A list of study inclusions and justifications for study exclusions are provided in Appendix 4.

Figure 1 QUOROM diagram



4.2 Number and type of studies included in the review of clinical effectiveness

4.2.1 Number and type of studies included in the review

Seven placebo-controlled RCTs were included in the clinical effectiveness review.^{14;15;19-26} Five of these were RCTs in patients with RRMS,^{14;15;19-23} while the remaining two trials evaluated the effect of disease-modifying therapies in patients with SPMS.²⁴⁻²⁶ Two head-to-head trials which compared IFN β -1b 8MIU²⁷ and IFN β -1a 44 μ g²⁸ with IFN β -1a 6MIU were also included in the review. Several publications were available for each trial. A summary of the studies included in the review is shown in Table 4.

Two of the studies evaluated GA 20mg; these were the European-Canadian trial for 9 months,¹⁴ the Copolymer-1 Multiple Sclerosis Study Group (C1MSSG) trial for 24 months¹⁹ (and up to 35 months for those patients who opted into the extension phase²⁰). The PRISMS trial compared two doses (22 μ g and 44 μ g) of subcutaneous IFN β -1a over a follow-up period of 24 months.²³ The Multiple Sclerosis Collaborative Research Group (MSCRG) trial studied intra-muscular IFN β -1a 6MIU (Avonex) over a duration of up to 24 months.²¹ The Interferon Beta Multiple Sclerosis Study Group (IFNBMSSG) trial of IFN β -1b 8MIU in RRMS¹⁵ followed patients up for a period of up to 5 years. The European Study Group^{24;25} and the North American Secondary Progressive Multiple Sclerosis trial (NASPMS)²⁶ studied IFN β -1b 8MIU in patients with SPMS, both for up to 36 months. Two trials included an additional intervention group in which patients were randomized to receive off-label doses of IFN β -1b;^{15;26} the results for these intervention arms are not included in this review.

With the exception of intra-muscular IFN β -1a 6MIU, which was administered by a health professional, injections were usually self-administered by the patient. Patients assigned to placebo received the same volume and frequency of injections but did not receive active medication. One trial²⁶ used two intervention groups, only one of which was relevant to this review; accordingly this trial also used two placebo conditions (50% randomized to a fixed 1.0mL of placebo solution, other half received a fraction of 1.0mL based on body surface area). However, results for the placebo groups were reported as an aggregate and are detailed here as such. The majority of trials specified that concomitant medication for treatment of relapses, usually steroids, and medication for influenza-like symptoms, were allowed throughout the trials. One trial specified that more than 3 courses of medication for relapses during 1 year would lead to withdrawal from the trial; this is likely to have resulted in bias, although the direction and magnitude of this is not clear.¹⁵

Table 4 Design and characteristics of included studies

Trial	Study design	Intervention 1	Intervention 2	Type of MS	Number randomized	Setting	Primary outcome
European-Canadian (2001) ¹⁴	Randomized placebo-controlled trial, double-blind	GA 20mg (Copaxone)	Placebo	RRMS	239 (GA 119 placebo 120)	Europe, Canada (29 centers)	MRI measures (relapse rate included as a tertiary outcome)
C1MSSG (1995, 1998) ^{19;20}	Randomized placebo-controlled trial, double-blind	GA 20mg (Copaxone)	Placebo	RRMS	251 (GA 125 placebo 126)	USA (11 centers)	Mean number of relapses
PRISMS (1998) ²³	Randomized placebo-controlled trial, double-blind	IFNβ-1a 22μg (Rebif)	IFNβ-1a 44μg (Rebif)	RRMS	560 (22μg 189 44μg 184 placebo 187)	Europe, Canada, Australia (22 centers)	Relapse rate
MSCRG (1996) ²¹	Randomized placebo-controlled trial, double-blind	IFNβ-1a 6MIU (Avonex)	Placebo	RRMS	301 (IFNβ-1a 158 placebo 143)	USA (4 centers)	Time to sustained disease progression
IFNBMSSG (1993, 1995) ^{15;22}	Randomized placebo-controlled trial, double-blind	IFNβ-1b 8MIU (Betaseron)	Off-label dose IFNβ-1b 1.6MIU Placebo	RRMS	372 (IFNβ-1b 124 placebo 123 off-label dose 125)	USA, Canada (11 centers)	Relapse rate, and proportion of patients relapse-free
European Study Group (1998, 2001) ^{24;25}	Randomized placebo-controlled trial, double-blind	IFNβ-1b 8MIU (Betaseron)	Placebo	SPMS	718 (IFNβ-1b 360 placebo 358)	Europe (32 centers)	Time to sustained disease progression
NASPMS (1999) ²⁶	Randomized placebo-controlled trial, double-blind	IFNβ-1b 8MIU (Betaseron)	Off-label dose IFNβ-1b 5MIU per m squared, body surface area Placebo	SPMS	939 (IFNβ-1b 317 placebo 308 off-label dose 314)	USA, Canada (35 centers)	Time to sustained disease progression
INCOMIN (2002) ²⁷	Randomized head-to-head trial, unblinded, no placebo	IFNβ-1b 8MIU (Betaseron)	IFNβ-1a 6MIU (Avonex)	RRMS	188 (IFNβ-1b 96 IFNβ-1a 92)	Italy (15 centers)	Proportion of patients relapse-free (and MRI measures)
EVIDENCE (2002) ²⁸	Randomized head-to-head trial, unblinded, no placebo	IFNβ-1a 44μg (Rebif)	IFNβ-1a 6MIU (Avonex)	RRMS	677 (44μg 339 6MIU 338)	Europe, Canada, USA (56 centers)	Proportion of patients relapse-free

Table key

C1MSSG	Copolymer-1 Multiple Sclerosis Study Group
PRISMS	Prevention of Relapses and Disability by Interferon β1a Subcutaneously in Multiple Sclerosis
MSCRG	Multiple Sclerosis Collaborative Research Group
IFNBMSSG	Interferon Beta Multiple Sclerosis Study Group
NASPMS	North American Secondary Progressive Multiple Sclerosis
INCOMIN	INdependent COMparison of INterferons
EVIDENCE	Evidence of Interferon Dose-response: European North American Comparative Efficacy
IFNβ	Interferon beta
GA	Glatiramer acetate

4.2.2 Number and type of studies excluded from the review

Eight studies that were not RCTs were excluded from the review. In addition, two publications of ongoing trials were also excluded from the review as detailed outcome data were not available at the time of writing. Off-label indications within trials, or off-label doses or administration routes were excluded from the review (n=8), as were trials which included only one on-label intervention without a placebo comparator arm (n=2). Four trials that were published only as abstracts/letters were also excluded from the clinical effectiveness review. Two trials were also excluded as the patients enrolled in the study had not been diagnosed with clinically definite MS. Six trials that did not report EDSS data were also excluded. Details of the excluded studies, and justifications for their exclusion, are reported in Appendix 4.

4.3 Assessment of trial quality

The seven randomized placebo-controlled trials included in the review^{14;15;19-26} were considered separately from the two head-to-head trials with IFN β -1a 6MIU^{27;28} due to their differing study designs. Specific issues surrounding randomization, blinding, comparability of baseline patient characteristics and statistical analysis methods for the included trials are detailed in Sections 4.3.1 and 4.3.2 below.

4.3.1 Quality of placebo-controlled trials

4.3.1.1 Randomization

The randomization strategy was adequate in six of the trials,^{14;20;21;23;24;26} and unclear in one trial.¹⁵ Concealment of treatment assignment was adequate in five trials,^{14;21;23;24;26} and unclear in two trials.^{15;20}

4.3.1.2 Blinding

In all seven placebo-controlled RCTs, patients, providers and outcome assessors were blinded to treatment group.^{14;15;20;21;23;24;26} However, blinding may have been compromised by adverse events in intervention groups; these were mainly injection-site reactions. In all seven trials, examining neurologists/physicians were separate from treating physicians. In some cases, outcome assessors were instructed not to discuss symptoms,^{14;20;21;24;26} and some trials explicitly stated that injection sites should be covered up during assessment.^{23;24;26} Questionnaires were conducted to assess blinding.^{15;21;24;26} The Phase III placebo-controlled trial of IFN β -1a 6MIU in the treatment of RRMS reported that 99% of outcome assessors did not know treatment assignment, 32.2% patients guessed correctly, and fewer treating physicians than patients guessed correctly.²¹ The placebo-controlled trial of IFN β -1b 8MIU in the treatment of RRMS reported that 80% of patients in the intervention group and 30% of patients in the placebo group guessed their treatment assignment correctly at the end of the study.¹⁵ In the European SPMS trial, 65.6% of IFN β -1b 8MIU patients guessed correctly and 54.3% of placebo patients guessed correctly; 56.2% and

48.4% of treating physicians guessed correctly for IFN β -1b 8MIU and placebo respectively, whereas outcome assessors guessed correctly for 20.8% of the IFN β -1b 8MIU group and 18.6% of the placebo group.²⁴ At the end of the North American SPMS trial, 26% of patients, 45% of treating physicians and 85% of evaluating physicians guessed “don’t know” when asked about their treatment assignment.²⁶

4.3.1.3 Baseline comparability of treatment groups

Where reported, the trials demonstrated similarities in terms of baseline EDSS, disease duration, patient age, relapse rate, gender and ethnicity between intervention and placebo groups.^{14;15;20;21;23;24;26}

4.3.1.4 Analytical approach within clinical trials

All seven trials had planned intention-to-treat analyses (ITT), although these were not conducted in all cases. ITT analyses were undertaken for some outcomes for some trials.^{14;23;24;26} For one trial, EDSS data were restricted to patients with sufficient time points of data, and other analyses were restricted by patients who accrued 104 weeks follow-up (approximately half those randomized).²¹ Analysis using data until drop-out was employed in one trial.¹⁵ In some cases, missing data were imputed using Last Observation Carried Forward (LOCF),^{14;20} or different analyses were conducted for EDSS, using calculations with lost to follow up as all progressed or all not progressed.²⁴ These imputation methods are unlikely to yield unbiased estimates of treatment effects.

4.3.2 Quality of head-to-head trials

4.3.2.1 Randomization

The generation of randomization sequences was adequate for both head-to-head trials included in the review.^{27;28} For both trials, treatment assignment was conducted centrally but study personnel and patients were unblinded.^{27;28}

4.3.2.2 Blinding

Patients and providers were not blinded due to different administration routes and frequency of medication. One study attempted to blind evaluating physicians by instructing them not to discuss symptoms and by covering injection sites; a questionnaire found that 52% of outcome assessors guessed treatment assignment correctly.²⁸

4.3.2.3 Baseline comparability of treatment groups

Within the trials there were reported similarities in baseline EDSS, disease duration, relapse rate during the previous two years and patient age between intervention and placebo groups.^{27;28}

4.3.2.4 Analytical approach within clinical trials

The analysis of both head-to-head trials^{27,28} was undertaken according to the ITT principle. For one trial, missing data were handled by retaining data until the point of being lost to follow-up, then an unfavorable outcome was assumed.²⁷ For the other trial, missing data concerning relapses were imputed using random number allocation based on the combined groups' proportion of patients not experiencing a relapse during the follow-up periods.²⁸

4.3.3 Study withdrawals

4.3.3.1 Study withdrawals in placebo-controlled trials

Study withdrawals may impact on trial results due to missing data; this may indicate problems with compliance due to medication or adverse events. Table 5 details the number of patients who withdrew from the placebo-controlled trials, and the number of patients who were lost to follow-up. Most trials reported a fairly low number of patients who dropped out due to adverse events.

Two hundred and fifteen patients completed 24 months on study within the Phase III trial of GA 20mg. Two hundred and three of these patients entered the 11 month extension study.²⁰ The Phase III trial of IFN β -1a 6MIU was ended prematurely; 73 patients in the intervention group and 56 patients in the placebo group did not complete 24 months of treatment.²¹ The North American IFN β -1b 8MIU trial was also terminated prematurely, although most of the intended data collection had been completed.²⁶ The Phase III trial of IFN β -1b 8MIU in RRMS commenced its fifth year with 58 patients in the intervention group and 56 in the placebo group, although only 5 patients were reported to have completed five years of treatment.¹⁵

Table 5 Study withdrawals for placebo-controlled trials, RRMS and SPMS

Trial	Number of patients randomized	Study duration (years)	Treatment group	Number of withdrawals (including lost to follow-up)	Number lost to follow-up	Number of withdrawals due to adverse events
Comi et al (2001) ¹⁴	239	0.75	GA 20mg	7	0	3
			Placebo	7	1	2
C1MSSG (1995, 1998) ^{19,20}	251	2	GA 20mg	19	NR	5
			Placebo	17	NR	1
PRISMS (1998) ²³	560	2	IFN β -1a 22 μ g	22	12	6
			IFN β -1a 44 μ g	19	5	9
			Placebo	17	10	2
MSCRG (1996) ²¹	301	2	IFN β -1a 6MIU	14	5 (unclear which group)	7
			Placebo	9		2
IFNBMSSG (1993, 1995) ^{15,22}	372	3	IFN β -1b 8MIU	25	2	10
			Placebo	23	2	1
European Study Group (1998, 2001) ^{24,25}	718	3	IFN β -1b 8MIU	103	40	47
			Placebo	117	48	15
NASPMS (1999) ²⁶	939	3	IFN β -1b 8MIU	79	22	30
			Placebo	69	11	32
Table key						
NR=not reported						

4.3.3.2 Study withdrawals for head-to-head trials

Table 6 shows the number of patients who withdrew due to adverse events or were lost to follow-up within the two head-to-head trials.^{27;28}

Table 6 Study withdrawals for head-to-head trials, RRMS

Trial	Number of patients randomized	Study duration (years)	Treatment group	Number of withdrawals (including lost to follow-up)	Number lost to follow-up	Number of withdrawals due to adverse events
INCOMIN (2002) ²⁷	188	2	IFN β -1b 8MIU	11	2	5
			IFN β -1a 6MIU	19	4	1
EVIDENCE (2002) ²⁸	677	0.92	IFN β -1a 44 μ g	25	0	16
			IFN β -1a 6MIU	21	1	14

4.4 Inclusion/exclusion criteria employed in trials

All trials included only patients with clinically definite MS. For trials in patients with RRMS, inclusion criteria specified that disease must have been diagnosed for a period of at least 1 year,^{14;15;21;23} or that the onset of the first relapse was at least 1 year prior to randomization.²⁰ Trials differentially specified at least one relapse in 2 years prior to randomization,¹⁴ at least two relapses in 2 years prior to randomization,^{15;20;23;27;28} or at least 2 documented relapses in the previous 3 years.²¹ Trials additionally specified that patients must be relapse-free 30 days prior to study entry,^{14;15;20;27} or relapse-free 2 months prior to study entry,²¹ with an EDSS score of 0-5,^{14;20;23} 0-5.5,^{15;28} or 1-3.5.^{21;27}

For trials in patients with SPMS, inclusion criteria were diagnosis of SPMS with deterioration for at least 6 months, following on from RRMS, with baseline EDSS of 3.0-6.5,^{24;26} and no relapses within either 30 days²⁴ or 60 days²⁶ of study entry. Trials included adult patients, generally aged 18 to 50 or 55, but with an upper age limit of 45 years in one trial,¹⁹ and 65 years in another trial.²⁶

4.5 Baseline characteristics of patients enrolled in included studies

Although treatment groups within trials had similar baseline characteristics, there were differences between trials in terms of baseline disease state and population demographics (See Table 7).

excluding relapses and unscheduled evaluations,²⁴ or confirmed at 2 visits across 6 months²⁶). The North American trial²⁶ had older patients with fewer relapses than the European Study Group trial.²⁴ Owing to this considerable heterogeneity between trial populations and outcome definitions, statistical meta-analysis to produce pooled estimates of efficacy for individual disease-modifying therapies was considered to be inappropriate.

4.7 Sustained disease progression outcomes

4.7.1 *Trial definitions of sustained disease progression*

Definitions of sustained disease progression differed between the included studies. Sustained progression was defined as an increase of at least 1.0 step on the EDSS,^{15;19;21-28} with some studies additionally stipulating a 0.5 step if baseline EDSS was 6 or above²⁴⁻²⁶ with confirmation that increased EDSS had been maintained for at least 3 months^{15;19;20;22-25;28} or 6 months.^{21;26;27}

4.7.2 *Sustained progression outcomes, placebo-controlled trials in RRMS*

4.7.2.1 *Sustained disease progression hazards, placebo-controlled trials in RRMS*

Table 8 presents the number of patients who experienced sustained disease progression within the placebo-controlled trials in RRMS which met the inclusion criteria for the review. Most of these study publications present analyses of time to sustained progression outcomes between treatment groups and placebo. However, none of the included studies in RRMS present the relative hazard ratio which relates to the difference between the empirical time to progression curves. Instead, estimates of the relative hazard ratios for disease progression between treatment groups have been estimated by the study authors using the methods described in Section 3.6.2.

relative hazard ratio for progression for IFN β -1b 8MIU versus placebo is 0.71. Based upon the 2-year results of the PRISMS trial,²³ the relative hazard ratios for progression for IFN β -1a 44 μ g and IFN β -1a 22 μ g versus placebo were estimated to be 0.60 and 0.72 respectively. It should be noted that the standard of reporting within the RCTs of disease-modifying therapy for the treatment of patients with RRMS is poor; none of the study publications report hazard ratios describing differences in time-to-event outcomes between treatment groups. Consequently, relative hazard ratios were estimated by the study authors based upon the assumption that progression-free survival is exponentially distributed (See Section 3.6.2). Standard errors and/or confidence intervals surrounding these relative hazard ratios were not available from published data.

4.7.2.2 Mean change in EDSS, placebo-controlled trials in RRMS

Table 9 presents reported mean changes in EDSS score between baseline and study end for the placebo-controlled trials in RRMS included in the review.

Table 9 Mean change in EDSS, placebo-controlled trials in RRMS

Study	Study duration (years)	Treatment group	Reported mean change in EDSS within study period	95% confidence interval	Significance (mean difference in EDSS change)
Comi et al (2001) ¹⁴	0.75	GA 20mg	0.02	NR	p=NS (statistical test unclear)
		Placebo	0.05	NR	
Johnson et al (1995) ¹⁹	2	GA 20mg	-0.05	+/-1.13 (SE)	p=0.023 (RMA of covariance)
		placebo	0.21	+/-0.99 (SE)	
Johnson et al (1998) ²⁰	2.92	20mg GA	-0.11	-0.31 to +0.10	p=0.02 (repeated measures of ANCOVA)
		placebo	0.34	+0.13 to +0.54	
MSCRG (1996) ²¹	2	IFN β -1a 6MIU	0.25 †	0.14 (SE)	p=0.02 (Mann-Whitney rank sum test)
		Placebo	0.74 †	0.16 (SE)	
IFNBMSSG (1993) ²²	3	IFN β -1b 8MIU	NR	NR	NR
		Placebo	NR	NR	
IFNBMSSG (1995) ¹⁵	5	IFN β -1b 8MIU	0.11 (median annual EDSS change)	NR	NR
		Placebo	0 (median annual EDSS change)	NR	
PRISMS (1998) ²³	2	IFN β -1a 44 μ g	0.24	1.1 (SD)	p<0.05 (ANOVA)
		IFN β -1a 22 μ g	0.23	1.3 (SD)	
		Placebo	0.48	1.3 (SD)	
Table key					
NR=not reported					
SE=Standard error					
SD=Standard deviation					
RMA=Repeated measures analysis					
†Results for IFN β -1a 6MIU vs. placebo relate to within-person change in EDSS from baseline to week 104 for those patients who completed 104 weeks in study.					

A non-significant difference in mean EDSS change from baseline was reported for GA 20mg and placebo at 9-months.¹⁴ However, a statistically significant improvement was reported for patients receiving GA 20mg compared to placebo at 24-months (p=0.023)¹⁹ and 35-months (p=0.02).²⁰

Significant differences in mean EDSS change from baseline to study end were reported in favor of IFN β -1a 6MIU (p=0.02),²¹ IFN β -1a 44 μ g,²³ and IFN β -1a 22 μ g,²³ compared to placebo (p<0.05).

4.7.2.3 Other EDSS-related outcomes, placebo-controlled trials of RRMS

The Phase III trial of GA 20mg versus placebo reported that more GA patients improved by 1+ EDSS step and more patients in the placebo group worsened by 1+ EDSS step; this difference was statistically significant at 24-months (p=0.024)¹⁹ and 35-months (p=0.001).²⁰ This trial also reported statistically significantly longer time to worsening by 1.5 EDSS steps in the GA treatment group than in placebo group.²⁰

The publication of the placebo-controlled trial of IFN β -1a 6MIU versus placebo²¹ reported the probability of onset of sustained progression during year 1 and year 2 (study end). Of those patients who remained in the study for at least 104 weeks, 21.8% of patients receiving placebo and 12.9% of patients receiving IFN β -1a 6MIU became treatment failures during year 1. During year 2, 14.7% of patients receiving placebo and 9.5% of patients receiving IFN β -1a 6MIU became treatment failures.²¹

Within the placebo-controlled trial of IFN β -1b 8MIU,¹⁵ time to progression was also evaluated separately for patients with a baseline EDSS<3.0 and those with a baseline EDSS \geq 3.0. No significant differences were found between treatment groups either within the lower EDSS group (p=0.535) or the higher EDSS group (p=0.087).¹⁵

The PRISMS study²³ reported that for patients with a high baseline EDSS score (>3.5), time to sustained progression was significantly longer for patients receiving IFN β -1a 44 μ g than for patients receiving placebo (p<0.05).

4.7.3 Disease progression outcomes for placebo-controlled trials in SPMS

4.7.3.1 Sustained disease progression hazards, placebo-controlled trials in SPMS

Table 10 presents the number of patients who experienced sustained disease progression within the placebo-controlled trials in SPMS included in the review. As with the data reported for patients with RRMS, relative hazard ratios were not available from the study publications; instead these were estimated by the study authors based upon methods described in Section 3.6.2.

confidence intervals surrounding these hazard ratios were not available from published trial reports.

4.7.3.2 Mean change in EDSS, placebo-controlled trials in SPMS

Table 11 presents mean changes in EDSS score between baseline and study end for the placebo-controlled trials in SPMS included in the review.

Table 11 Mean change in EDSS, placebo-controlled trials in SPMS

Study	Study duration (years)	Treatment group	Reported mean change in EDSS within study period	95% confidence interval	Significance (mean difference in EDSS change)
European Study Group (1998) ²⁴	2	IFNβ-1b 8MIU	0.47	NR	0.0299 (Extended Mantel-Haenszel test)
		Placebo	0.60	NR	
European Study Group (2001) ²⁵	3	IFNβ-1b 8MIU	0.47	NR	0.003 (Extended Mantel-Haenszel test)
		Placebo	0.69	NR	
NASPMS (1999) ²⁶	3	IFNβ-1b 8MIU	0.53	NR	p>0.05 (ANOVA)
		Placebo	0.62	NR	
Table key					
NR=not reported					

Significant differences in mean change in EDSS from baseline to study end in favor of IFNβ-1b 8MIU over placebo were reported at 2 years²⁴ and 3 years²⁵ within the European SPMS trial (p<0.03). These significant differences in mean EDSS change from baseline to study end were however not observed within the NASPMS trial (p>0.05).²⁶

4.7.3.3 Other EDSS-related outcomes, placebo-controlled trials in SPMS

The European placebo-controlled trial of IFNβ-1b 8MIU²⁴ reported that time to sustained progression was longer for those patients receiving IFNβ-1b 8MIU (days to event 40% quantile 893, lower 95% confidence interval limit 726, upper limit not estimated) than for those receiving placebo (days to event 40% quantile 549, 95% confidence interval 463-642). A piecewise logistic regression analysis suggested an odds ratio for progression of 0.65 (95% confidence interval 0.52-0.83); this analysis did not suggest the existence of interactions between the treatment and other variables included in the statistical model, thus suggesting a homogeneous treatment effect over time.²⁴ The authors also report probabilities of progressing by each 3-month time-point; this analysis suggested that the difference in progression between treatment and placebo became statistically significant after 12 months (p=0.0003) and was maintained for the remainder of the study duration. Logistic regression suggested that the odds of progression for treatment versus placebo was 0.63 (95% confidence interval 0.46-0.85). The study authors also reported a significant difference between the treatment groups in terms of time to becoming wheelchair-bound in favor of IFNβ-1b 8MIU (p=0.0133). Significant differences were also observed in terms

of the probability of becoming wheelchair-bound in favor of the IFN β -1b 8MIU group (year 1 p=0.0129, year 2 p=0.0094, year 3 p=0.0133).²⁴

The final 3-year results of the European trial²⁵ also reported significant differences in favor of IFN β -1b 8MIU over placebo in terms of the proportion of patients with sustained progression for patients with >2 relapses 2 years prior to study entry (p=0.04); the proportion of patients with sustained progression for patients with a >1 point change in EDSS 2 years before entering the study (p=0.02); the proportion of patients with sustained progression for patients who experienced >1 point change in EDSS or relapse 2 years before entering the study (p=0.03); the proportion of patients with sustained progression for patients with a duration of MS greater than 11.9 years (p=0.01); the proportion of patients with evidence of progressive deterioration greater than 3 years (p=0.01); the proportion of patients older than 42 years of age (p=0.04); the proportion of patients who were female (p=0.03); time to becoming wheelchair-bound (year 1 p=0.014, year 2 p=0.010, year 3 p=0.005); and the proportion of patients becoming wheelchair-bound (p=0.007). Other non-significant differences in proportions of patients who experienced disease progression favored the IFN β -1b 8MIU group rather than placebo.

Within the NASPMS trial,²⁶ further analyses were undertaken using 3-month rather than 6-month confirmation of sustained disease progression; this analysis did not lead to a significant difference in time to progression between the study groups. Other exploratory *post-hoc* endpoints included analyses of time to progression by center, the potential effect of pre-study exacerbations on apparent progression, potential effects of concomitant medications, time of study entry, US versus Canadian patients, as well as other prognostic factors. None of these analyses produced a significant difference in time to progression between the IFN β -1b 8MIU and placebo groups.²⁶

4.7.4 Disease progression outcomes for head-to-head trials in RRMS

4.7.4.1 Sustained disease progression hazards, head-to-head trials in RRMS

Table 12 presents the number of patients reported to have experienced sustained disease progression within the included head-to-head trials in patients with RRMS. Hazard ratios for time to sustained progression were reported for the EVIDENCE trial²⁸ but were not reported for the INCOMIN trial;²⁷ the hazard ratio for the INCOMIN trial was estimated by the study authors using the methods reported in Section 3.6.2.

4.7.5 Modified relative hazard ratios for progression based upon the placebo-controlled trials and head-to-head trials in RRMS

Table 13 presents the results of the random effects mixed treatment comparisons, which synthesize evidence on sustained disease progression from the placebo-controlled trials^{15;21;23} and the head-to-head trials^{27;28} in patients with RRMS. The methods used to generate these estimates are presented in Section 3.6.4, and the WinBUGS model syntax and data are detailed in Appendix 3. It should be noted that none of the studies which met the inclusion criteria for this review reported head-to-head trials of GA 20mg against either IFN β -1a or IFN β -1b. As there are differences between the trials in terms of patient characteristics and definitions of sustained progression, the results of this analysis should be interpreted tentatively. Owing to these potential heterogeneities between the studies, the impact of the mixed treatment comparisons on marginal estimates of cost-effectiveness is considered as a sensitivity analysis (See Section 6.3.2).

Table 13 Progression hazard results estimated using mixed treatment comparisons of placebo-controlled and head-to-head trials in RRMS

Study	Study duration (years)	Treatment group	N	Observed number of events	WinBUGS predicted number of events	WINBUGS estimated relative hazard (SE)
MSCRG (1996) ²¹	2	IFN β -1a 6MIU	158	34.60	40.36	0.79 (0.12)
		Placebo	143	49.91	44.73	
IFNBMSSG (1995) ¹⁵	5	IFN β -1b 8MIU	122	43	39.34	0.52 (0.09)
		Placebo	122	56	64.65	
PRISMS (1998) ²³	2	IFN β -1a 44 μ g	184	47.84 [†]	48.68	0.70 (0.11)
		Placebo	187	72.93 [†]	67.38	
INCOMIN (2002) ²⁷	2	IFN β -1b 8MIU	96	13	17.07	0.67 (0.14)
		IFN β -1a 6MIU	92	28	23.68	
EVIDENCE (2002) ²⁸	0.92	IFN β -1a 44 μ g	339	43	41.92	0.89 (0.15)
		IFN β -1a 6MIU	338	49	47.31	
Table key						
<i>SE=standard error</i>						
<i>† Estimated from Kaplan-Meier progression-free survival curves</i>						

It can be seen from the results of the mixed treatment comparisons that the incorporation of evidence from the two head-to-head trials^{27;28} has a substantial impact upon the relative hazard ratios for EDSS progression estimated from the placebo-controlled trials alone (estimated hazard ratios for independent and mixed treatment comparisons – IFN β -1a 6MIU 0.58 versus 0.79, IFN β -1b 8MIU 0.71 versus 0.52, and IFN β -1a 44 μ g 0.60 versus 0.70 for the placebo-controlled and WinBUGS synthesis respectively). Table 13 suggests that the WinBUGS synthesis places considerably more weight on the progression results reported for the EVIDENCE trial²⁸ due to its large sample size. The WinBUGS analysis indicates that the head-to-head estimates of the effectiveness of IFN β -1a 44 μ g are broadly consistent with the placebo-controlled data, while the WinBUGS estimates of the effectiveness of IFN β -1b 8MIU appear to be more favorable than those obtained from the analysis of the placebo-controlled trial data. By contrast, the head-to-head trials suggest that the placebo-controlled data for 6MIU IFN β -1a result in an overestimate of its true effect on progression. However, as noted above, the presence of known and unknown

heterogeneities between the trials mean that the results of this analysis should be approached with caution.

4.7.6 Discussion of EDSS outcomes

For placebo-controlled trials in patients with RRMS, significant improvements in terms of time to sustained EDSS progression were reported for IFN β -1a 6MIU,²¹ IFN β -1a 44 μ g,²³ and IFN β -1a 22 μ g.²³ For head-to-head trials in patients with RRMS, a significant improvement in time to sustained EDSS progression was reported for IFN β -1b 8MIU compared to IFN β -1a 6MIU.²⁷ IFN β -1a 44 μ g did not significantly delay time to disease progression compared to IFN β -1a 6MIU.²⁸

For placebo-controlled trials in patients with SPMS, significant improvements in terms of time to sustained EDSS progression and mean EDSS change from baseline to study end were reported for IFN β -1b 8MIU within the European study,^{24;25} but not within the NASPMS trial.²⁶ This contrasting finding between the two trials may be explained by the different baseline populations and/or by the different definitions of progression.

Where no significant differences were reported between intervention and comparator groups, trends suggested beneficial rather than detrimental effects of interventions over the comparator. However, a lack of statistical significance means that firm conclusions cannot be drawn. Heterogeneity in terms of the definition of disease progression between studies means that comparisons of results between studies are highly problematic, especially with regard to differing times of confirmation of EDSS scores. Progression confirmed at 3-months may constitute prolonged relapse that would not be confirmed at 6-months, as indicated where both definitions were measured within a trial.²⁵

A further point of note is that the analysis of the GA 20mg outcomes^{19;20} accepted results reflecting improvement on the EDSS, while the other included studies do not. However, observed improvements on the EDSS are generally interpreted as indicating the erroneous measurement of disease progression at the previous examination.²³ Mean change in EDSS is also not a valid measure where the baseline population has a range of EDSS scores, as EDSS is an ordinal scale, thus changes at one end of the scale do not match measured change at the other end of the scale. The nature of the EDSS scale was sometimes taken into account through the use of differing definitions of disease progression at or above an EDSS score of 6.0.²⁴⁻²⁶

Differing distributions of patients across disease states at study baseline may also compromise direct comparisons between trials. It is unclear whether the interventions have a differential effect on specific subgroups of patients. The Phase III trial of subcutaneous IFN β -1a reported a more

pronounced treatment effect in high baseline EDSS (over 3.5) patients.²³ Patients with relapses or increases in EDSS two years prior to study entry were relatively affected more by IFN β -1b 8MIU in both SPMS trials.^{25;26}

It should also be noted that the reporting of progression outcomes within the included studies is limited. While most of the studies compared time-to-event data using the log-rank test, the corresponding relative hazard ratios and uncertainty surrounding these ratios were only reported within the EVIDENCE trial.²⁸ For all other included studies, we estimated relative hazard ratios between treatment groups based upon the assumption that the effect of treatment is independent of time (See Section 3.6.2).

It should also be noted that the study periods used within the trials to evaluate the impact of therapy on disease progression were short. Longer follow-up periods may provide a better indication of the true effect of the disease-modifying therapies on sustained disease progression.

4.8 Relapse outcomes

4.8.1. *Trial definitions of relapse*

Relapse was defined as the appearance or reappearance of one or more neurologic symptoms, immediately preceded by a relatively stable or improving neurologic state of at least 30 days, persisting for either at least 48 hours,^{14;20;21;26} or persisting for at least 24 hours^{15;23;24} with some studies additionally stipulating symptoms consistent with at least 0.5 step increase on EDSS,^{14;20;21} or in the absence of fever.^{15;24;26} The definitions of relapse were similar for studies in RRMS and SPMS.

4.8.2. *Relapse outcomes, placebo-controlled trials in RRMS*

4.8.2.1 *Relapse rates, placebo-controlled trials in RRMS*

Table 14 presents the mean number of relapses experienced by patients within each treatment group as reported in the publications of placebo-controlled trials in patients with RRMS.

Table 14 Relapse rates, placebo-controlled trials in RRMS

Study	Study duration (years)	Treatment group	N	Mean relapse rate	Annualized relapse rate (95% confidence interval)	Relative relapse rate (95% confidence interval)	Significance (difference in relapse rate)
Comi et al (2001) ¹⁴	0.75	GA 20mg	119	0.51	0.81 (NR)	0.67 (NR)†	p=0.012 (statistical test unclear)
		Placebo	120	0.76	1.21 (NR)		
Johnson et al (1995) ¹⁹	2	GA 20mg	125	1.19	0.59 (NR)	0.70 (NR)†	p=0.007 (ANCOVA)
		placebo	126	1.68	0.84 (NR)		
Johnson et al (1998) ²⁰	2.92	20mg GA	125	1.34	0.58 (1.06-1.63)	0.72 (NR)†	p=0.002 (ANCOVA)
		placebo	126	1.98	0.81 (1.70-2.26)		
MSCRG (1996) ²¹	2	IFNβ-1a 6MIU	158	NR	0.67 (NR)	0.82 (NR)†	p=0.04 (Mann-Whitney rank sum test)
		placebo	143	NR	0.82 (NR)		
IFNBMSSG (1993) ²²	3	IFNβ-1b 8MIU	124	NR	0.84 (NR)	0.66 (NR)†	p=0.0001 (ANOVA)
		placebo	123	NR	1.27 (NR)		
IFNBMSSG (1995) ¹⁵	5	IFNβ-1b 8MIU	122	NR	0.78 (0.70 to 0.88)	0.70 (NR)†	p=0.0006 (ANOVA)
		placebo	122	NR	1.12 (1.02 to 1.23)		
PRISMS (1998) ²³	2	IFNβ-1a 44μg	184	1.73	NR (NR)	0.68 (95% confidence interval for relative relapse rate reduction 0.21-0.44)	p<0.005 (generalized linear model)
		IFNβ-1a 22μg	189	1.82	NR (NR)		
		Placebo	187	2.56	NR (NR)	0.71 (95% confidence interval for relative relapse rate reduction 0.14-0.39)	p<0.005 (generalized linear model)

Table key
NR=Not reported
†Confidence intervals surrounding relative relapse rates were not reported in the trial publications and have been estimated by the study authors using methods described in Section 3.6.2.

All of the placebo-controlled RCTs in RRMS included in the review reported mean annualized relapse rates with the exception of the PRISMS trial which reported only unadjusted mean relapse rates.²³ Significant improvements in relapse rate were reported for all interventions compared against placebo.^{14;15;19-23} Significant differences in relapse rates in favor of the intervention groups were reported within both interim analyses and final analyses for IFNβ-1b 8MIU and GA 20mg. For IFNβ-1a 6MIU, significant improvements were reported in favor of the intervention both when all patients were included (p=0.04), and when the analysis was restricted to patients with 104 weeks data (p=0.002).²¹

The mean relative relapse rate for GA 20mg compared to placebo was estimated to be 0.67¹⁴ to 0.72.²⁰ The estimated mean relative relapse rate for IFN β -1a 6MIU versus placebo was estimated to be 0.82.²¹ The mean relative relapse rate for IFN β -1b 8MIU versus placebo was estimated to be 0.66 when the 3-year outcomes were used,²² and 0.70 when the 5-year outcomes were used.¹⁵ The mean relative relapse rate for IFN β -1a 44 μ g versus placebo was estimated to be 0.68 while the relative relapse rate for IFN β -1a 22 μ g versus placebo was estimated to be 0.71.²³ As with many of the other clinical efficacy outcomes, insufficient information was provided within the study publications to estimate standard errors or confidence intervals surrounding these relative relapse rates.

4.8.2.2 Other relapse outcomes, placebo-controlled trials in RRMS

Median time to first relapse from baseline for GA 20mg was reported to be 287 days; for placebo median time to first relapse was reported to be 198 days, although this difference was not statistically significant ($p=0.097$).¹⁹ This same estimate was reported at 35 months.²⁰ 42 patients (33.6%) receiving GA 20mg and 34 patients (27.0%) remained relapse-free throughout the initial 2-year study duration ($p=0.098$). The therapeutic effect of GA 20mg on relapse was reported to be most pronounced in patients with EDSS scores between 0 and 2.0 at baseline. At 35-months, 33.6% of patients receiving GA 20mg and 24.6% of patients receiving placebo were reported to be relapse-free ($p=0.035$). The 35-month extension study report stated that patients receiving placebo were significantly more likely to suffer multiple relapses ($p=0.008$).²⁰ The study by Comi et al¹⁴ however did not report a significant difference in terms of the proportion of relapse-free patients for the GA 20mg and placebo groups ($p=0.175$).

The placebo-controlled trial of IFN β -1a 6MIU reported that patients receiving the active treatment were significantly less likely to experience multiple relapses ($p=0.03$), although this analysis included only those patients who had spent at least 104 weeks on the study.²¹ The annualized relapse rate for this subset of patients who completed 104 weeks on study was reported to be 0.61 for IFN β -1a 6MIU and 0.90 for placebo; this difference in relapse rate was reported to be statistically significant ($p=0.002$).²¹ A non-significant difference in median time to first relapse was reported in favor of IFN β -1a 6MIU (median time to relapse IFN β -1a 6MIU 36.1 weeks, placebo 47.3 weeks, $p=0.034$).²¹

The 3-year analysis of the placebo-controlled trial of IFN β -1b 8MIU in RRMS reported a significant difference in time to first relapse in favor of IFN β -1b 8MIU after 2 years of treatment (median time to first relapse IFN β -1b 8MIU 295 days, placebo 153 days, $p=0.015$) which continued after 3 years of treatment (median time to first relapse IFN β -1b 8MIU 264 days, placebo 147 days, $p=0.028$).²² At 2-years, a significant reduction in moderate/severe relapses was reported for IFN β -1b 8MIU compared to placebo ($p=0.002$).²² This was reported to have led

8MIU, mean relapse rate=0.16, standard deviation=0.34; for placebo, mean relapse rate=0.28, standard deviation =0.51).

4.8.3.2 Other relapse outcomes, placebo-controlled trials in SPMS

The 2-year publication of the European SPMS trial²⁴ reported a significant prolongation of time to first relapse for patients receiving IFNβ-1b 8MIU (median time to first relapse IFNβ-1b 8MIU 644 days, placebo 403 days, p=0.0030). The proportion of patients with moderate or severe relapses was also lower in the intervention group than the placebo group (43.6% patients on IFNβ-1b 8MIU, 53.1% patients on placebo, p=0.0083).²⁴ In the North American SPMS study, fewer moderate and severe relapses were reported for patients receiving IFNβ-1b than for patients receiving placebo (p-value not reported).²⁶

4.8.4 Relapse outcomes, head-to-head trials in RRMS

4.8.4.1 Relapse rates, head-to-head trials in RRMS

Table 16 presents relapse rates for the two head-to-head trials included in the review.

Table 16 Relapse rates, head-to-head trials in RRMS

Study	Study duration (years)	Treatment group	N	Mean relapse rate	Annualized relapse rate (95% confidence interval)	Relative relapse rate (95% confidence interval)	Significance (difference in relapse rates)
INCOMIN (2002) ²⁷	2	IFNβ-1b 8MIU	96	NR	0.5 (SE=0.7)	0.71 (standardized mean difference =-0.31 (-0.60 to -0.02))	p=0.03 (statistical test unclear)
		IFNβ-1a 6MIU	92	NR	0.7 (SE=0.9)		
EVIDENCE (2002) ²⁸	0.46	IFNβ-1a 44μg	339	0.29	NR (NR)	0.73 (NR)†	p=0.022 (Poisson regression model)
		IFNβ-1a 6MIU	338	0.4	NR (NR)		
	0.92	IFNβ-1a 44μg	339	0.54	NR (NR)	0.84 (NR)†	p=0.0093 (Poisson regression model)
		IFNβ-1a 6MIU	338	0.64	NR (NR)		
Table key							
NR=Not reported							
SE=Standard error							
†Confidence intervals were not reported in the trial publications and have been estimated by the study authors using methods described in Section 3.6.2.							

The INCOMIN trial publication²⁷ reported a statistically significant reduction in annualized relapse rate in favor of IFNβ-1b 8MIU compared to IFNβ-1a 6MIU (standardized mean difference 0.31, 95% confidence interval -0.60 to -0.02, p=0.03); this led to a relative relapse rate of 0.71. The EVIDENCE trial publication reported a statistically significant reduction in mean relapse rate for IFNβ-1a 44μg compared to IFNβ-1a 6MIU at both 24-weeks (relative relapse rate=0.73, p=0.022) and 48-weeks (relative relapse rate=0.84, p=0.00093).²⁸

4.8.4.2 Other relapse outcomes, head-to-head trials in RRMS

The EVIDENCE trial publication reported that over the initial 24 weeks of treatment, 254 (75%) patients receiving IFN β -1a 44 μ g and 214 (63%) patients receiving IFN β -1a 6MIU remained relapse-free ($p=0.0005$). Over 48 weeks of treatment, 209 patients (62%) receiving IFN β -1a 44 μ g and 177 patients (52%) receiving IFN β -1a 6MIU remained relapse-free ($p=0.009$). A statistically significant improvement in time to first relapse was reported in favor of patients receiving IFN β -1a 44 μ g (hazard ratio=0.70, $p=0.003$).²⁸ The proportion of mild, moderate and severe relapses were similar in each treatment group.

The INCOMIN trial reported a statistically significant difference in terms of the proportion of patients free from relapse over the study duration in favor of patients receiving IFN β -1b 8MIU (IFN β -1b 8MIU 33%, IFN β -1a 6MIU 51%, relative risk=0.76, 95% confidence interval 0.59-0.99, $p=0.03$).²⁷

4.8.5 Modified relative relapse rates based upon placebo-controlled trials and head-to-head trials in RRMS

Table 17 shows the results of the fixed effects mixed treatment comparison WinBUGS model, which synthesizes evidence on relapse rates from the placebo-controlled trials^{15;21;23} and the head-to-head trials^{27;28} in patients with RRMS. As noted earlier, none of the identified studies which met the inclusion criteria for this review reported head-to-head trials of GA 20mg against IFN β . As with the mixed treatment comparisons for sustained disease progression, caution is advised in interpreting the results of this WinBUGS analysis due to both known and unknown heterogeneities between the trials. The impact of these modified relapse rates on marginal estimates of cost-effectiveness is considered as a sensitivity analysis (See Section 6.3.2).

Table 17 Relative relapse rates estimated using mixed treatment comparisons of placebo-controlled and head-to-head trials in RRMS

Study	Treatment group	N	Observed mean number of relapses per patient	WinBUGS predicted relative relapse rate (standard error)
PRISMS (1998) ²³	IFN β -1a 44 μ g	184	1.73	0.68 (0.05)
	Placebo	187	2.56	
IFNBSG (1995) ¹⁵	IFN β -1b 8MIU	124	0.78	0.66 (0.07)
	Placebo	123	1.12	
MSCRG (1995) ²¹	IFN β -1a 6MIU	158	0.67	0.83 (0.07)
	Placebo	143	0.82	
INCOMIN (2002) ²⁷	IFN β -1b 8MIU	96	0.50	0.80 (0.10)
	IFN β -1a 6MIU	92	0.70	
EVIDENCE (2002) ²⁸	IFN β -1a 44 μ g	339	0.54	0.83 (0.07)
	IFN β -1a 6MIU	338	0.64	

The results of the mixed treatment comparisons suggest that after adjusting for relapse rates estimated using data from the two head-to-head trials,^{27;28} the estimated relapse rates for each

intervention compared to placebo are very similar to those estimates obtained from the placebo-controlled RCTs (IFN β -1a 6MIU = 0.82 vs. 0.83, IFN β -1a 44 μ g = 0.68 vs. 0.68, IFN β -1b 8MIU = 0.70 vs. 0.66 for placebo-controlled trial estimates and WinBUGS estimates respectively).

4.8.6. Discussion of relapse outcomes

Statistically significant reductions in relapse rates were shown for all of the interventions compared to placebo.^{14;15;19-23} For other secondary endpoints,^{24;26} differences failing to reach significance showed trends which favored the intervention over placebo. For the two included trials of IFN β -1b 8MIU in the treatment of patients with SPMS, significant reductions in relapse rates were demonstrated in favor of the intervention.^{24;26} The two head-to-head trials^{27;28} reported statistically significant improvements in relapse rates for IFN β -1b 8MIU and IFN β -1a 44 μ g when compared against IFN β -1a 6MIU.

4.9 Adverse events and quality of life

Injection site reactions were reported for all interventions; these included pain, itching, swelling and redness. Injection site reactions were more often observed in subcutaneous rather than intramuscular injections, and were more common for patients receiving IFN β -1a and IFN β -1b than GA. Influenza-like symptoms were reported as fairly frequent in interferon treatment (approximately half of the treatment groups^{15;21;23}) and were less common with GA treatment. There was rarer occurrence of fever, myalgia,^{15;21;23} fatigue, and headache.^{21;23} There were no significant differences between IFN β and placebo groups in depression or suicide.^{15;21;23} There was a transient self-limiting systemic reaction including flushing, chest pain and anxiety in some patients treated with GA (approximately 10%), but there was no major toxicity associated with GA.^{14;20}

HRQoL data were reported only for the European Phase III trial of IFN β -1b 8MIU in the treatment of patients with SPMS,²⁹ using the Sickness Impact Profile and a scale designed specifically for the study. The study showed a trend for delayed deterioration in patient-reported HRQoL, although this only reached statistical significance for a subscale (physical scale) of the Sickness Impact Profile intermittently (at 6-, 12- and 36-months, but not at other 6-monthly measurements).

4.10 Conclusions on clinical effectiveness

Trials demonstrated significant benefits for all interventions on relapses, shown by lower relapse rates, fewer patients experiencing relapse during study periods and delayed time to first relapse. All interventions significantly reduced relapse rates when compared with placebo. IFN β -1b 8MIU (Betaseron) and IFN β -1a 44 μ g (Rebif) both reduced relapse rates significantly more than IFN β -1a 6MIU (Avonex).^{27;28} Based upon the analysis of the placebo-controlled RCTs, the relative

reduction in relapse rate was estimated to be around 30% for IFN β -1a 44 μ g (Rebif), IFN β -1a 22 μ g (Rebif), IFN β -1b 8MIU (Betaseron) and GA 20mg (Copaxone) when compared against placebo. The relative reduction in relapse rate was estimated to be approximately 18% for IFN β -1a 6MIU (Avonex) when compared against placebo. When evidence from the head-to-head trials was included in the analysis, mean reductions in relapse rates for the interferons versus placebo appeared to be similar to those obtained from the placebo-controlled trials.

Evidence concerning sustained disease progression was equivocal. For placebo-controlled trials in patients with RRMS, statistically significant improvements in time to sustained disease progression were reported for IFN β -1a (Rebif) at both 44 μ g and 22 μ g doses, and for IFN β -1a 6MIU (Avonex) compared against placebo. Hazard ratios were not reported within these trials but were estimated by the study authors (See Section 3.6.2). The relative hazard ratio for progression for GA 20mg (Copaxone) versus placebo was estimated to be in the range 0.76 to 0.86. The estimated relative hazard ratio for progression for IFN β -1a 6MIU (Avonex) versus placebo was estimated to be 0.58; however, analysis of this trial was not undertaken according to the ITT principle. The relative hazard ratio for progression for IFN β -1b 8MIU (Betaseron) versus placebo was estimated to be 0.71. The estimated relative hazard ratios for progression for IFN β -1a 44 μ g and IFN β -1a 22 μ g (Rebif) versus placebo were estimated to be 0.60 and 0.72 respectively. Estimates of uncertainty surrounding these hazard ratios were not available (or calculable) from the data reported within the study publications. There were also significantly lessened increases in mean EDSS score for GA 20mg (Copaxone), IFN β -1a 22 μ g (Rebif), IFN β -1a 44 μ g (Rebif), IFN β -1a 6MIU (Avonex) and IFN β -1b 8MIU (Betaseron), however these data may reflect prolonged relapses rather than sustained progression.

For head-to-head trials in patients with RRMS, a significant difference in terms of time to sustained EDSS progression was reported for IFN β -1b 8MIU (Betaseron) compared to IFN β -1a 6MIU (Avonex) within the INCOMIN study.²⁷ Within the EVIDENCE trial, IFN β -1a 44 μ g (Rebif) did not significantly delay time to disease progression compared to IFN β -1a 6MIU (Avonex).²⁸ The mixed treatment comparisons model which synthesizes evidence from the head-to-head trials and placebo-controlled trials suggests that both IFN β -1a 44 μ g (Rebif) and IFN β -1b 8MIU (Betaseron) appear to be more effective in slowing progression than IFN β -1a 6MIU (Avonex), however, these results should be approached with caution.

For the treatment of SPMS, one study reported a significant improvement in time to disease progression for IFN β -1b 8MIU (Betaseron); a non-significant difference was reported for the other SPMS trial. The hazard ratio for IFN β -1b 8MIU versus placebo for SPMS was estimated to be between 0.72 and 0.93.

Study periods for considering disease progression were short; further research over longer experimental study durations would be invaluable. Differences in study designs, baseline populations, follow-up periods, primary endpoints, and definitions of relapse and progression endpoints make comparisons of efficacy between studies highly problematic. Missing data were not imputed in the same way across trials, and not all trials conducted ITT analyses, therefore the presence of bias is likely. Many of the trials recruited only a relatively small number of patients. Where subgroup analyses were conducted, there was a suggestion that disease status at baseline was a factor in the effectiveness of the intervention; consequently further research with planned subgroup analyses or comparisons of specific subgroups of patients would be useful in elucidating this potential interaction. Additional head-to-head trials may also indicate the relative effectiveness of GA and IFN β .

Injection-site reactions were common; these present a problem for blinding in trials of GA and IFN β . Other treatment related adverse events included influenza-like symptoms. Serious adverse events were rare, and there was no evidence to indicate that depression or suicide were linked to any of the interventions. There were very few drop-outs due to adverse events, suggesting that the interventions were generally well tolerated.

It should be re-iterated that the standard of reporting of outcomes within the included studies was generally limited, with very few studies reporting hazard ratios for time-to-event data, or more generally, confidence intervals or standard errors surrounding mean effectiveness estimates.

4.11 Summary of clinical effectiveness data used to populate the health economic model

Table 18 summarises the estimates of treatment effectiveness used to populate the base case health economic model. Owing to the limited reporting of relative hazard ratios for disease progression and relative relapse rates (and uncertainty surrounding these statistics), standard errors surrounding mean treatment effectiveness estimates were derived from a process of consultation with clinical experts, the National Institute for Clinical Excellence (NICE) and its stakeholders during the 2001 appraisal of these products. These same standard errors are used within the US analysis. It is noteworthy that these standard errors allow for *more* uncertainty than was observed in the re-analysis of the patient-level trial data; this additional uncertainty is particularly appropriate due to the gap between the clinical trial populations and the US MS Medicare population.

Table 18 Summary of treatment effectiveness estimates assumed within the base case health economic analysis

Intervention	Progression		Relapse	
	Relative hazard ratio for progression in RRMS (SE)	Relative hazard ratio for progression in SPMS (SE)	Relative relapse rate in RRMS (SE)	Relative relapse rate in SPMS (SE)
GA 20mg	0.86 (0.23)	1.00 (n/a)	0.70 (0.11)	1.00 (n/a)
IFNβ-1a 6MIU	0.58 (0.19)	1.00 (n/a)	0.82 (0.13)	1.00 (n/a)
IFNβ-1b 8MIU	0.71 (0.18)	0.72 (0.18)	0.70 (0.09)	0.69 (0.09)
IFNβ-1a 22μg	0.72 (0.19)	1.00 (n/a)	0.71 (0.08)	1.00 (n/a)
IFNβ-1a 44μg	0.60 (0.19)	1.00 (n/a)	0.68 (0.08)	1.00 (n/a)
Table key				
<i>SE=standard error</i>				

For the purpose of the base case health economic analysis, relative progression hazards and relative relapse rates for IFNβ-1a 6MIU, IFNβ-1b 8MIU, IFNβ-1a 22μg, and IFNβ-1a 44μg were derived from the final analyses of the pivotal clinical trials for these products.^{15;21;23} As noted throughout this chapter, the placebo-controlled trial of IFNβ-1a 6MIU was stopped early and many reported analyses were not based upon the ITT principle.²¹ While this is flawed, this remains the most substantial public domain estimate of the relative efficacy of IFNβ-1a 6MIU compared to placebo, and importantly is the estimate of effect upon which the product is licensed. For GA 20mg, effectiveness estimates were drawn from the 2-year analysis.¹⁹ The extension study outcomes for GA 20mg were not used within the model due to the use of the Last Observation Carried Forward rule to impute missing EDSS data; this imputation approach is likely to result in biased estimates of treatment effectiveness. For SPMS, the most favorable hazard ratio for IFNβ-1b 8MIU versus placebo was used.

Table 19 summarises the estimates of treatment effectiveness used within the mixed treatment comparison model which is presented as a sensitivity analysis. Standard errors surrounding mean effectiveness estimates are shown in parentheses. In this instance, standard errors surrounding mean effectiveness estimates were generated using the WinBUGS software.

Table 19 Summary of treatment effectiveness data assumed within the mixed treatment comparison sensitivity analysis

Intervention	Progression		Relapse	
	Relative hazard ratio for progression in RRMS (SE)	Relative hazard ratio for progression in SPMS (SE)	Relative relapse rate in RRMS (SE)	Relative relapse rate in SPMS (SE)
GA 20mg	0.86 (0.23) [†]	1.00 (n/a)	0.70 (0.11) [†]	1.00 (n/a)
IFNβ-1a 6MIU	0.79 (0.12)	1.00 (n/a)	0.83 (0.07)	1.00 (n/a)
IFNβ-1b 8MIU	0.52 (0.09)	0.72 (0.18) [†]	0.66 (0.07)	0.69 (0.09) [†]
IFNβ-1a 22μg	0.72 (0.19) [†]	1.00 (n/a)	0.71 (0.08) [†]	1.00 (n/a)
IFNβ-1a 44μg	0.70 (0.11)	1.00 (n/a)	0.68 (0.05)	1.00 (n/a)
Table key				
<i>[†]Based on placebo-controlled trial estimates due to absence of head-to-head trial comparisons</i>				

5.0 Health economic methods

5.1 Introduction

This chapter details the structure and data sources used within the ScHARR MS cost-effectiveness model. The model was originally developed for use in the appraisal of IFN β and GA on behalf of NICE in England and Wales. The original cost-effectiveness model has been developed to incorporate a lifetime horizon for the evaluation of the costs and effects of alternative treatment options, the incorporation of US-specific data on baseline disease severity,^{30;31} updated estimates of treatment effectiveness derived from the review of clinical effectiveness^{14;15;19-28} (See Chapter 4), updated estimates of HRQoL by EDSS state from a cohort of patients within Nova Scotia,³² and US-specific data on the costs of disease-modifying therapies and the costs of MS care.³⁰

5.2 Existing evidence on the health economics of IFN β and GA

There have been a number of attempts to estimate the cost-effectiveness of the disease-modifying therapies for MS. Previous estimates of cost-effectiveness range from in excess of a million pounds per QALY gained (>US \$1,800,000) to cost-saving (See “Key Definitions”).³³⁻³⁷ A review of these models undertaken on behalf of NICE identified significant flaws in all existing analyses.³⁸ The major problems identified in the review were:

- Failure to model the natural history of the disease as the comparator to treatment;
- Failure to incorporate mortality in long-term treatment models;
- Failure to model the transition to SPMS from RRMS;
- Failure to model the impact of treatment-related adverse events on cost-effectiveness;
- Failure to incorporate treatment drop-outs into the model;
- Linear extrapolation of short-term data;
- Inappropriate time horizons;
- Implausible assumptions regarding the impact of relapse on HRQoL;
- Inadequate analysis of uncertainty around model parameter values.

In March 2001, NICE commissioned a consortium based at the University of Sheffield to produce a cost-effectiveness analysis to directly address the problems associated with the existing cost-effectiveness analyses of the disease-modifying therapies for MS. The health economic model presented within this report is based upon the original model developed on behalf of NICE, with the inclusion of additional US-specific cost and utility data, and alternative model assumptions which are more appropriate to the cost-effectiveness analysis of disease-modifying therapies in the US Medicare MS population. The structure of the model and a description of the evidence used to populate the model have been reported previously.³⁹⁻⁴¹

5.3 Perspective of the analysis

The original model developed as part of the 2001 NICE appraisal evaluated the direct costs and effects of disease-modifying therapies from the perspective of the NHS and Personal Social Services (See “Key Definitions”). The model developed on behalf of the CMS evaluates direct costs and effects associated with the disease-modifying therapies from the perspective of the US health care payer.

5.4 Structure of health economic model

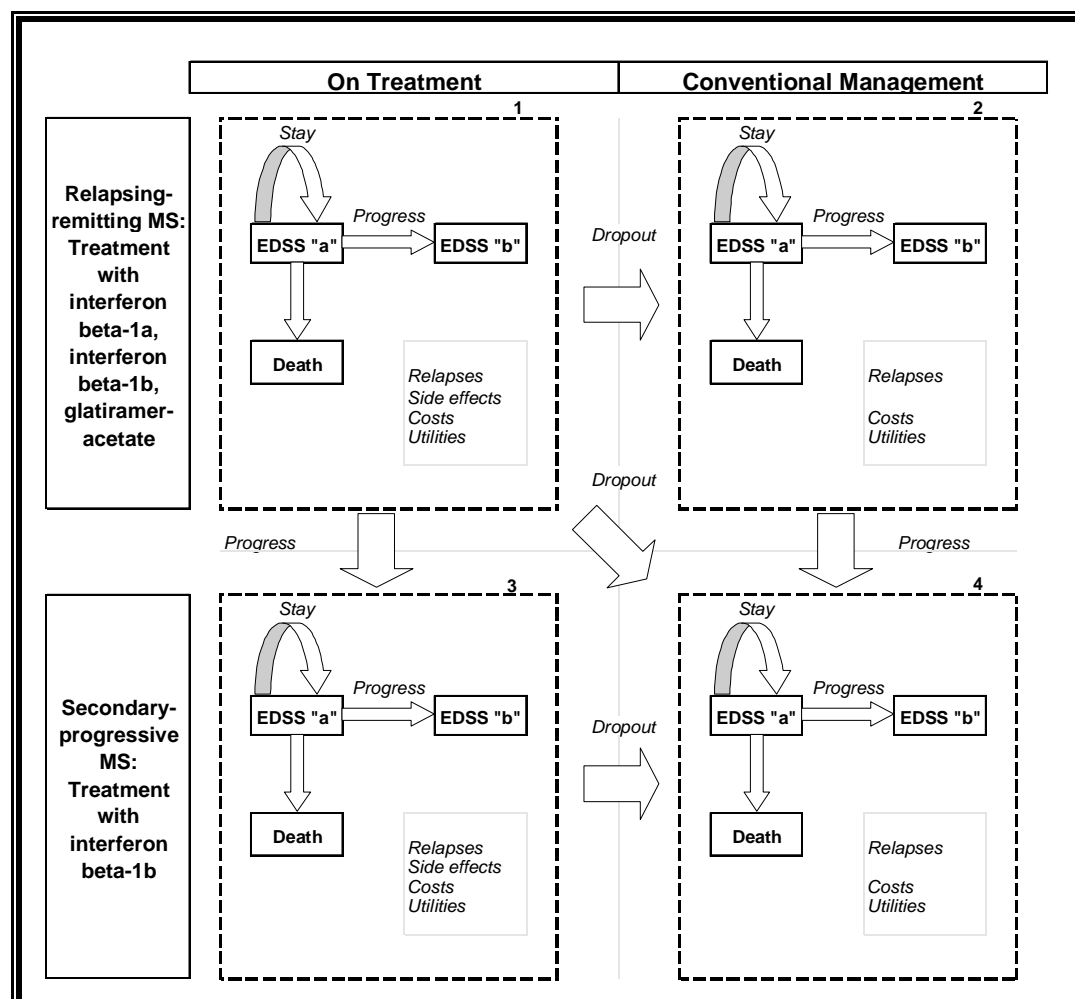
5.4.1 Overview of cost-effectiveness model

The ScHARR MS cost-effectiveness model simulates the natural history of MS using the state transition methodology, modeling individual EDSS states from 0 through 10 in RRMS, and from EDSS 2.0 through 10 in SPMS.⁷ The model operates on a cohort basis rather than at the level of the individual patient. Each health state in the model relates to an individual score on the EDSS. Patients enter the model at 51 years of age, based upon data obtained from the Sonya Slifka MS dataset (See Section 5.6.8.2).^{30;31} Disease course, cost and utilities with and without treatment are assessed over the remaining lifetime of the cohort. All patients enter the model and with rare exception, experience progressive disability as measured on the EDSS according to the transition probabilities derived from the London Ontario Cohort natural history dataset (See Section 5.6.1).⁴² Disease progression within the model is assumed to be time-independent, although these probabilities are adjusted as the risk of death increases. Improvements on the EDSS are assumed not to be possible within this model. Figure 2 shows the possible transitions between model health states during each cycle. During any given model cycle, patients can remain in their current health state, progress one or more EDSS states, transit to a SPMS health state, discontinue therapy, or die. Patients who discontinue therapy subsequently progress according to the natural history transition rates.⁴² Importantly, patients are assumed to retain the benefits of treatment achieved up to the point of stopping therapy.

Costs are associated with spending one cycle in each individual EDSS health state. Each EDSS state is assigned a specific utility score which describes the mean quality of life associated with that degree of disability. The “on treatment” and “conventional management” (best supportive care) cohorts progress through the EDSS states at different rates, as determined by instantaneous hazard rates estimated using current clinical effectiveness evidence (See Chapter 4), and therefore accrue different cost and quality of life profiles.

Useful guides concerning the application of the state transition modeling methodology in the economic evaluation of healthcare technologies are available from Sonnenberg and Beck⁴³ and Briggs and Sculpher.⁴⁴

Figure 2 Schematic of ScHARR MS cost-effectiveness model



5.4.2 Key model assumptions

The assumptions made in constructing the model favor the novel therapies within the analysis.

- Patients enter the model aged 51 years, as reflected in the Sonya Slifka dataset.^{30;31}
- The initial distribution of patients across the EDSS is based upon the ADL distribution of all patients in the Sonya Slifka dataset³⁰ at the time of the analysis (mapped to the EDSS, as described in Section 5.6.8.2).
- While some patients with primary progressive disease may receive disease-modifying therapy in practice in the US, there is a paucity of evidence relating to the clinical efficacy of any disease-modifying therapy within this patient population. Consequently, the analysis is restricted only to patients with RRMS and SPMS disease courses.
- The simulation is continued until all patients within the model are absorbed into the death state.
- Transitions within the model are assumed to be progressive only. For example, a patient in EDSS 4.5 in the current model cycle could not regress back to EDSS 4.0 during a

subsequent model cycle. Fluctuations due to exacerbations and subsequent stabilization are superimposed upon the underlying EDSS progression model. It has been argued that MS may not in fact be a solely progressive disease, but rather that fluctuations on the EDSS may be possible. It is unclear whether these observed improvements in disability are 'noise' or whether short-term improvements in disability actually occur. Noise in the data could stem from the misspecification of initial or subsequent disability, or variations in patients' attitudes to their underlying functioning levels over time. However, such improvements have not been observed in the long-term natural history data from Ebers and colleagues,⁴² and were not considered within the majority of the clinical trials (See Chapter 4). Current evidence suggests that the appearance of shifts towards improved disability is likely to be attributable to the poor intra- and inter-rater reliability of the EDSS.

- A 'retained effect' of treatment on both progression and relapse beyond the duration of the trials included in the clinical effectiveness review is modeled. Any patient who discontinues therapy subsequently progresses according to natural history rates but retains any previously accrued benefits at no additional cost of therapy.
- Due to the paucity of evidence concerning the long-term efficacy of any of these therapies, the effects of treatment are assumed to be fixed and do not deteriorate or increase over time.
- The annual risk of 'all-cause' mortality for the MS cohort is assumed to be the same as a normal healthy population.
- Patients may continue to receive disease-modifying therapy until they drop off therapy or until death. Assumptions concerning the impact of treatment cessation upon progression to EDSS 7.0 are explored within the sensitivity analysis.

5.5 Health economic comparisons undertaken within the cost-effectiveness analysis

It is conventional practice in health economic evaluation to compare the costs and effects of health interventions incrementally, whereby interventions are ranked in order of effectiveness, and cost-effectiveness ratios are calculated for non-dominated treatment options. However, correlations between the efficacies of the range of IFN β s and GA are unknown and have not been fully evaluated within clinical trials; if one therapy is effective, it is possible that the other therapies are also effective. The review of placebo-controlled trials presented in Chapter 4 highlighted considerable heterogeneities between the studies in terms of the populations of patients at baseline and in terms of the definition of progression outcomes. In order to incorporate such correlations between the efficacies of the disease-modifying therapies within the model, it would be necessary to handle the set of treatment efficacies as a multivariate normal distribution, and to incorporate an

uncertain covariance matrix into the model. Sampling of this multivariate distribution would then be facilitated by sequentially sampling a series of standardized normal distributions, and linearly transforming these samples using the Cholesky square root of the covariance matrix. This situation is further complicated by the necessity to also sample the covariance matrix in order to capture the uncertainty in the correlations between treatments.

Owing to these difficulties, the base case health economic model reports the marginal cost-effectiveness of each of the disease-modifying therapies as compared against best supportive care. This analysis is based exclusively on effectiveness data sourced from the placebo-controlled clinical trials (effectiveness estimates are reported in Table 18). The synthesis of placebo-controlled trials and head-to-head trials of IFN β provides some information concerning the comparative efficacy of some of these therapies (IFN β -1a and IFN β -1b), although evidence concerning the comparative efficacy of GA and IFN β is not available. A secondary analysis of cost-effectiveness based upon modified treatment effectiveness estimates using both the placebo-controlled and head-to-head trials is therefore presented as a sensitivity analysis (effectiveness estimates are reported in Table 19).

5.6 Data sources used within the SchARR MS cost-effectiveness model

A list of all model parameters and their characteristics is detailed in Appendix 5.

5.6.1. Disease progression

Natural history disease progression rates for patients with RRMS and SPMS were derived from a large 25-year patient-level cohort study undertaken in London, Ontario, Canada.⁴² The clinical database was derived from a population-based sample of more than 1,000 patients with essentially untreated MS in whom accrual ended in 1984. This dataset was used to estimate instantaneous hazard rates for transition along the EDSS for patients receiving best supportive care within the model. Patients who were not eligible for treatment according to label indications and the Association of British Neurologists (ABN) guidelines⁴⁵ were excluded from the analysis. A summary of these criteria are shown in Box 1.

Box 1 ABN Eligibility Criteria for Treatment with IFN β in Multiple Sclerosis

- Patients must be able to walk independently
- Patients must have previously had at least two major relapses
- Patients must be aged at least 18 years
- Patients must not have any contraindications

Data on Disability Status Scale (DSS) score, a precursor to the EDSS,⁴⁶ for each patient for each year up to 25 years were used to estimate the probability of progressing from one DSS state to another (See Appendix 6). As natural history data were available at the level of the individual patient, it was possible to remove data for those patients with primary progressive or benign MS, to estimate separate transition rates for RRMS and SPMS, and to estimate the probability of progressing from RRMS to SPMS for individual DSS states.

Instantaneous hazard rates for progression without disease-modifying therapy were calculated from the Ontario dataset⁴² using the formula below. As transitions were based upon instantaneous hazard rates, multiple EDSS progressions are possible during any model cycle, for example it is possible for a patient to progress from EDSS 3.0 to EDSS 5.0 within a single model cycle.

$$I_i = \frac{\text{Number of people leaving state}_i}{\sum_{j=1}^n \text{duration in state}_i}$$

Pooled placebo arm data from the existing RCTs of IFN β and GA were used to map from the DSS transition matrix to an EDSS state transition matrix. Each DSS state was assumed to consist of two EDSS states i.e. DSS 4 was assumed to consist of EDSS 4.0 and EDSS 4.5. The proportion of time in each of the two EDSS states observed in the pooled placebo matrix was used to partition the DSS dwell time, while keeping the expected total dwell time in the two EDSS states equal to the observed dwell time in the natural history data. This method is reported in detail in the earlier study report.³⁹

5.6.2 Treatment effects on progression

The effect of treatment on disease progression was modeled using relative hazard rates as compared to no treatment, assuming independent treatment effects. The methods for estimating relative hazard rates from the included RCTs are described in Section 3.6. It should be noted that the model assumes that the treatment effect parameters estimated using data from the placebo-controlled trials are maintained over the entire model time horizon, although most trials were followed up over a period of between 2 and 3 years. As noted in Section 4.10, the true impact of

these therapies on disease progression and relapse over the course of the disease is subject to considerable uncertainty.

The relative hazard ratios for disease progression were modeled by applying the relative hazard ratio for disease progression for each treatment versus placebo to the underlying natural history disease progression rate. Therefore each annual transition probability is multiplied by the relative hazard ratio for disease progression.

5.6.3 Natural history relapses

The London Ontario cohort study⁴² collected very little data on relapses beyond the first 2 years of follow-up, and was thus of limited use in modeling the incidence of relapse for patients under usual clinical management. A recent review of the natural history and epidemiology of MS by Richards and colleagues⁴ on behalf of the National Health Service Health Technology Assessment Programme identified six studies that reported natural history relapse rates. Of those identified studies which reported relapse rates in terms of time since diagnosis, there was significant variation in terms of both the published relapse rates as well as the length of follow-up. The two most recent studies reported annual relapse rates for a period of only 2-3 years.^{47;48} The next most recent study, which was undertaken by Patzold and colleagues,⁴⁹ reported relapse rates over a considerably longer duration (a total of 19 years). Up to 7 years, these data are reported annually, and every two years thereafter. The Patzold data⁴⁹ as reported in Richards and colleagues⁴ were used as the basis for estimating the natural history of relapses. These data are presented in Table 20.

Table 20 Natural history relapse rates from time of diagnosis⁴⁹

Time (years)	Estimated annual natural history relapse rate†
1	1.85
2	1.1
3	1
4	0.85
5	0.65
7	0.75
9	0.25
11	0.6
13	0.28
15	0.3
19	0.2
Table key	
†Standard errors/95% confidence intervals not reported within study publication	

Evidence on the incidence of relapse clearly suggests that the frequency of relapse is time-dependent rather than EDSS-dependent.⁴ In order to identify the expected number of relapses for each EDSS state over time, the state transition matrix within the SchARR model was used to identify the number of years with disease for patients in each EDSS state. For each EDSS state, the distribution for disease duration was estimated. These proportions were then applied as weights to the time-dependent relapse rates taken from the Patzold study.⁴⁹ The products of each of these calculations were then summed to give the expected number of relapses in each state, independent of time since diagnosis.

5.6.4 Treatment effects on relapse

The effect of treatment on the mean number of relapses experienced was modeled using the relative relapse rate estimated from the clinical trials (See Section 4.8). As noted in Chapter 4, the magnitude of this effect differs between the therapies. This means that as patients transit from one health state to another, the mean number of relapses they experience during the next year depends on which health state they move to. As patients progress to SPMS, they incur fewer relapses as progression becomes their primary symptom. The occurrence of a relapse results in both an additional cost to be added to the patient's current total costs, and a decrement in utility to be subtracted from the patient's current total utility score.

Relative relapse rates while receiving each of the disease-modifying therapies compared to best supportive care were taken directly from relapse data reported within the clinical trials (See Section 4.8). These relative relapse rates were then applied directly to the Patzold natural history relapse data⁴⁹ to estimate the expected number of relapses for each treatment group over the model time horizon.

5.6.5 Treatment-related adverse events

It is well established that all four disease-modifying therapies are associated with adverse events of some type. The IFN β s are recognized as having injection site reactions and influenza-like symptoms, while GA is reported to have different, but equally self-resolving adverse events (for example hot flashes). The incidence of mild adverse events is high for all these products, although all four disease-modifying therapies are described as being well tolerated. The model assumes that the incidence of adverse events is as reported in the trials. However, the model also assumes that these adverse events occur only during the first year of treatment. While treatment-related adverse events are clinically mild, work by Prosser³⁷ suggests that they are associated with a small but significant reduction in HRQoL. Mean adverse event utility decrements estimated by Prosser³⁷ were used within the SchARR MS cost-effectiveness model.

5.6.6 Drop-outs

Published evidence suggests a drop-out rate for all therapies of up to around 30% during the trial durations, however, the evidence on the distribution of drop-outs over time is unclear. Evidence from actual usage provided by two of the companies suggested a slightly lower level of drop-outs, however this was still over 20%.³⁹ The drop-out estimates reported within the trials and the company-generated drop-out estimates cannot be considered robust. The trial data include drop-outs for different reasons, some of which are protocol-related, while the data on therapy cessation in clinical practice provide no information on the reasons for drop-out. As there appears to be little significant difference between treatment-related adverse events or their ease of use, it seems sensible to treat the therapies equally in this regard. Therefore the model assumes a drop-out rate of 10% in each of the first 2-years, followed by a 3% drop-out rate for each subsequent year on therapy. The underlying assumption is that there are a group of people who drop out early due to treatment-related adverse events (10%) and a second group of people who drop out later because either they or their doctors feel that the therapy is not helping (10%). After that point the model assumes a long-term attrition each year in use consistent with treatment compliance seen for the treatment of many chronic conditions.

The model does *not* include the possibility of switching between disease-modifying therapies. The model assumes that patients who drop out retain the benefits of treatment previously accrued but subsequently progress according to the natural history data. The effect of this assumption is that patients who cease therapy retain these benefits without incurring additional costs.

5.6.7 HRQoL

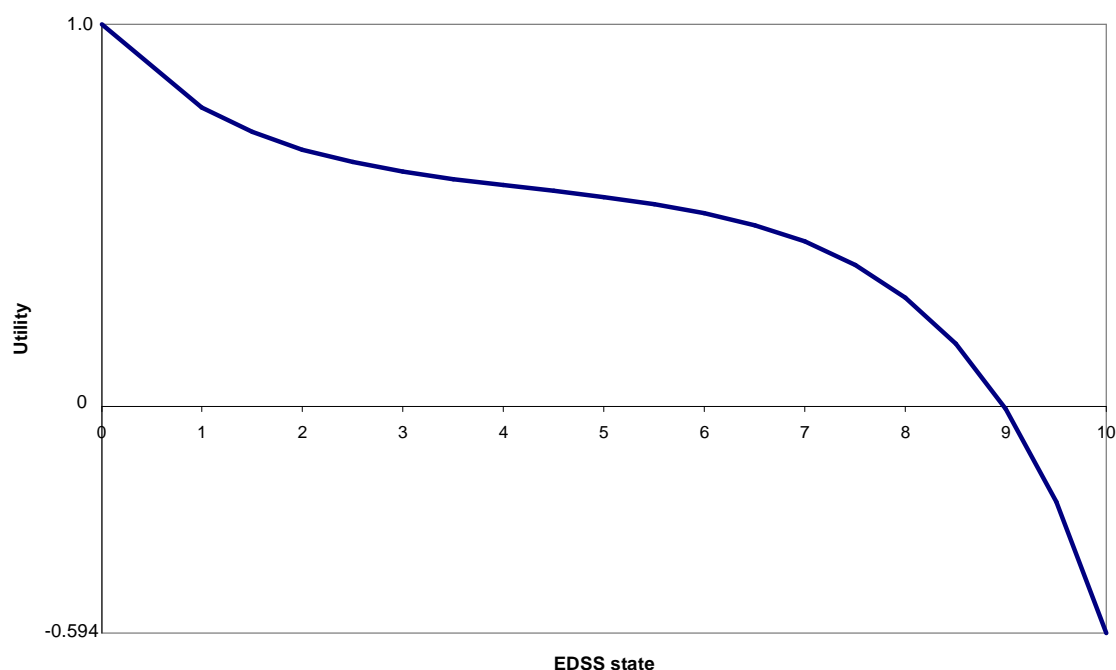
HRQoL was modeled by assigning a utility score to each health state within the model. The total number of QALYs gained for each treatment option was calculated as the survival in each health state over the time horizon multiplied by the EDSS-specific utility associated with each health state. The negative quality of life impacts associated with relapses and side effects are represented by disutilities which were subtracted from the total number of QALYs gained.

5.6.7.1 Modeling the relationship between EDSS and HRQoL

There is increasing evidence of robust relationships between an individual's EDSS and both the costs of managing their condition and the quality of life associated with that health state.^{50;51} However, the model developed on behalf of NICE did not recognize the correlation between EDSS states for either utilities or costs (i.e. as costs increase, utilities decrease in a systematic pattern reflecting the change in the underlying clinical condition). As part of a further study commissioned by the NHS Research and Development Programme, a revised cost-effectiveness model was developed using publicly available information on the relationship between the EDSS, costs of care and HRQoL.⁴¹ Mean utilities were replaced by a functional form based upon the

patient's underlying EDSS score. The specification of the functions drew from data available in the literature,^{33;50-52} our own experience of analyzing cost and quality of life data in MS, as well as our knowledge of methodological issues around cost and quality of life assessment in other chronic disabling conditions. On the latter point, we adjusted the functional form to allow for potential informative censoring in those patients who are severely disabled individuals, i.e. patients in EDSS states 8.0 and above. It was assumed that individuals at EDSS 9.5 would fulfil the criteria for the worst health state in the EQ-5D classification, thus the utility for this state was assumed to be -0.594. Previous experience in analyzing this type of data established that the shape of the function was consistently of the form illustrated in Figure 3.

Figure 3 Illustration of assumed relationship between EDSS and health status



While there is some variation in the published quality of life estimates, this is not substantial. In order to specify the function it was necessary to identify the quality of life value at the point for EDSS 6.0, where the function commences its sharper downward trajectory. Parkin et al³³ report a value of 0.49 for this state. As stated above, we were concerned that as disability increases, empirical estimates will be affected by the ability or willingness of disabled individuals to complete quality of life questionnaires. We therefore believe that mean values are likely to be overestimates of the true mean value for the health state. We therefore adjusted the Parkin³³ estimate slightly downwards to reflect this belief. We set the quality of life for EDSS 6.0 at a value of 0.47.

5.6.7.2 Incorporating recent empirical data with publicly available utility estimates

The utility data used within the model developed on behalf of the CMS retained the form of the utility function described in Section 5.6.7.1, and incorporated further unpublished Canadian utility data made available to the SchARR.³² It should be noted that equivalent EDSS-specific utility data were not available for the US Medicare MS population.

Data on HRQoL, as measured using the Health Utilities Index Mark 3 (HUI3) were obtained from a study of 813 patients with clinically definite MS in Nova Scotia, Canada. Patients within the study had at least one HUI3 assessment within 16 days of a clinical examination where EDSS score was collected. Of these patients, 784 fully completed the HUI3 instrument and were included in the analysis; the remainder did not complete all dimensions of the HUI3 questionnaire which precluded the calculation of an index utility score. These data are reported in Table 21.

Table 21 Relationship between utility score and EDSS³²

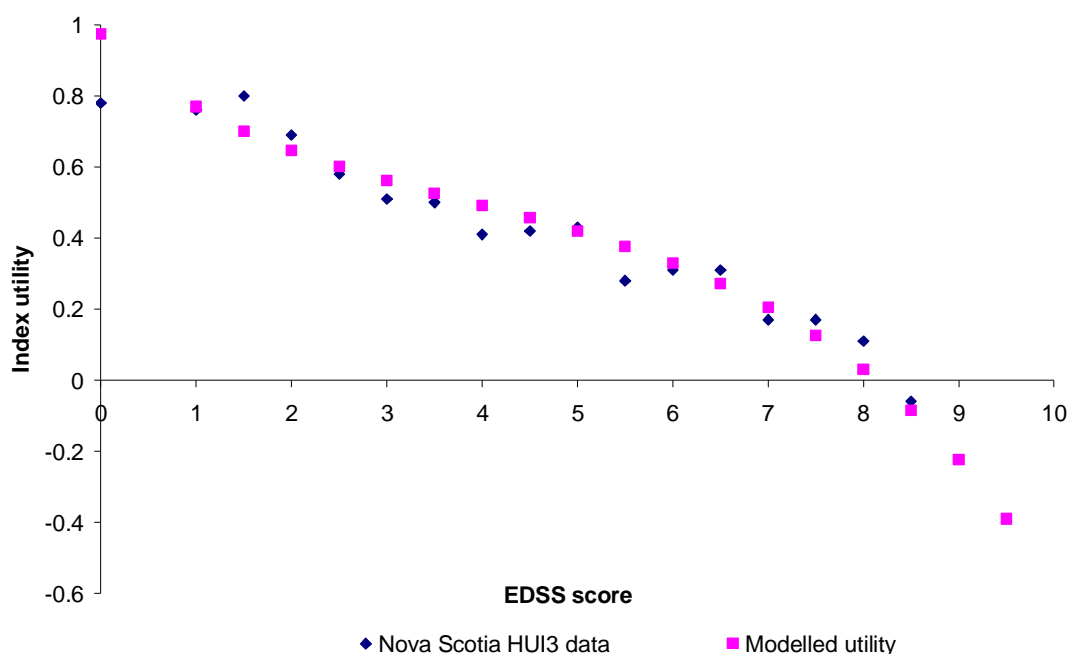
EDSS score	No. patients completing EDSS (n=813)	No. patients completing HUI3 (n=784)	Mean HUI3 score	Standard deviation	Lower 95% confidence interval	Upper 95% confidence interval
0	20	20	0.78	0.31	0.64	0.92
1.0	73	70	0.76	0.25	0.70	0.81
1.5	57	56	0.80	0.22	0.74	0.86
2.0	153	149	0.69	0.25	0.65	0.73
2.5	99	93	0.58	0.31	0.51	0.64
3.0	84	82	0.51	0.27	0.45	0.57
3.5	78	72	0.50	0.29	0.43	0.57
4.0	24	24	0.41	0.33	0.27	0.55
4.5	14	14	0.42	0.20	0.31	0.53
5.0	22	21	0.43	0.35	0.27	0.59
5.5	12	12	0.28	0.27	0.11	0.45
6.0	58	57	0.31	0.30	0.23	0.39
6.5	48	47	0.31	0.23	0.24	0.38
7.0	21	21	0.17	0.23	0.06	0.27
7.5	17	16	0.17	0.23	0.05	0.30
8.0	18	17	0.11	0.19	0.01	0.21
8.5	15	13	-0.06	0.18	-0.18	0.05

It is interesting to note that the anticipated missing data for those patients who are most severely disabled is not evident within the dataset.

The parameters to the EDSS-utility function were fitted to the empirical Nova Scotia data using least squares minimization techniques. The results of this fitting process are shown in Figure 4. Uncertainty surrounding mean utility scores was estimated using Monte Carlo sampling techniques; the 95% confidence intervals of these simulated EDSS-specific utility scores were then fitted to the confidence intervals observed within the Nova Scotia cohort using least squares minimization. Figure 4 demonstrates that the shape of the functional form provides a close approximation of the empirical Nova Scotia utility dataset. The two main benefits of this approach are that the uncertainty surrounding mean utility scores can be modeled appropriately through the

incorporation of correlations between EDSS points, and that the missing data for EDSS 9.0 and EDSS 9.5 can be predicted.

Figure 4 EDSS-utility function using data from Nova Scotia cohort



5.6.8 Costs

5.6.8.1 Acquisition and administration costs of disease-modifying therapies

All costs were valued in 2005 US dollars. Annual drug and administration costs for the disease-modifying therapies were obtained from the CMS at manufacturers recommended doses. As with other Part B drug therapies, physician-administered IFN β -1a (Avonex) is reimbursed on the basis of its Average Sales Price (ASP) plus 6%. In July 2005, the allowed drug charge for 11mcg IFN β -1a (Avonex) was \$90,368. This equates to an annual cost per patient of \$14,097. An additional administration cost of \$22 per weekly dose for physician services and clinic fees related to the administration of the drug was calculated from 2004 Medicare claims data for patients receiving IFN β -1a (Avonex); this cost was inflated to 2005 prices using Medicare fee schedules. This administration cost was added to the allowed charges for the drug, which resulted in a total allowed charge of \$15,257.64 for physician-administered IFN β -1a (Avonex). As Medicare reimburses 80% of allowed charges under Part B, the cost of physician-administered IFN β -1a 6MIU (Avonex) to Medicare is estimated to be \$12,438.57. It should be noted that self-administered IFN β -1a (Avonex) is not covered under the remit of Part B, but is included as a treatment option within the analysis.

In order to estimate drug costs for self-administered IFN β -1a 6MIU (Avonex), IFN β -1a (Rebif), IFN β -1b 8MIU (Betaseron) and GA 20mg (Copaxone) which were covered under the MRDD, drug claims data from the demonstration were analyzed by the CMS to estimate a mean daily cost for each treatment. The patient cost sharing arrangements for the drugs covered by the demonstration were complicated and were designed to reflect the arrangements for the Part D drug benefit scheme which commenced in January 2006. The mean Medicare payment under the demonstration for MS therapies was estimated to be 77.65%; this was applied to the four therapies covered under the demonstration. The annual costs of self-administered IFN β -1a 6MIU (Avonex), IFN β -1a (Rebif), IFN β -1b 8MIU (Betaseron) and GA 20mg (Copaxone) to Medicare were estimated to be \$11,533.59, \$14,803.78, \$13,084.67, and \$12,310.18 respectively. The same cost of 22 μ g and 44 μ g Rebif was used in the analysis, as Serono offer price parity for both doses of the drug. Administration costs are not relevant for the self-administered medications covered under the demonstration. A summary of these cost estimates are shown in Table 22.

Table 22 Estimated costs of disease-modifying therapies to Medicare

	IFN β -1a 6MIU (Avonex) physician- administered	IFN β -1a MIU (Avonex) self- administered	IFN β -1a 22 μ g (Rebif)	IFN β -1a 44 μ g (Rebif)	IFN β -1b 8MIU (Betaseron)	GA 20mg (Copaxone)
Daily cost	\$90.37	\$40.69	\$49.28	\$49.28	\$43.55	\$40.98
Annual cost	\$14,097.41	\$14,853.30	\$17,985.62	\$17,985.62	\$15,897.01	\$14,956.06
Annual administration charges	\$1,160.64	n/a	n/a	n/a	n/a	n/a
Medicare reimbursement	\$12,438.57	\$11,533.59	\$13,965.83	\$13,965.83	\$12,344.03	\$11,613.38

5.6.8.2 Estimating the direct costs of MS care

The Sonya Slifka dataset

Recent published evidence relating to the relationship between the degree of MS disability and costs of care in the US is limited.⁵⁰ A review of existing evidence concerning the costs associated with different levels of MS disability as measured by the EDSS, identified only two US studies^{53;54} which had estimated the costs associated with specific levels of disability. The review reported by Patwardhan and colleagues⁵⁰ suggested that rise in cost is positively correlated with the level of MS disability.

Patwardhan and colleagues⁵⁰ noted that few studies could be identified in which the investigators performed a comprehensive assessment of the entire spectrum of the cost associated with MS. Owing to the absence of good quality MS cost evidence available in the public domain, EDSS-specific costs of care were obtained from an analysis of data collected within the Sonya Slifka database commissioned specifically for use in this cost-effectiveness analysis.⁵⁵ This work was undertaken by ABT Associates, Inc. in conjunction with the CMS and the authors of this report in

order to produce an analytical file of costs associated with health care resources used by participants in the Sonya Slifka Longitudinal Multiple Sclerosis Study with the explicit purpose of estimating the cost-effectiveness of the disease-modifying therapies. This sample included only those patients who represented the Medicare population receiving these therapies.

At the time of the analysis, the Sonya Slifka study sample consisted of 2,156 people with MS who have been shown to be representative of the MS population in the US.³¹ Owing to the lack of empirical evidence on the efficacy of the disease-modifying therapies in PPMS and RPMS, the SchARR MS cost-effectiveness model estimates the cost-effectiveness of these therapies in the management of RRMS and SPMS. Therefore, a subset of the wider Sonya Slifka dataset was identified and selected for use in the model. This sample consisted of those patients with a RRMS or SPMS disease course who had received the same disease-modifying therapy for at least six months prior to the 12-month interview, as agreed with the CMS.

A large range of resource use items were collected and contained within the Slifka dataset;³⁰ these included direct costs such as hospitalizations, outpatient and emergency room visits, treatments, laboratory tests, health care professional visits and medications. The dataset also included indirect cost items (See “Key Definitions”) such as lost productivity costs for patients and carers, and the costs of consuming social and health services, for example home modifications. Where possible, unit costs were assigned to resource use items based upon the reimbursed amount. For the majority of resource utilization items, unit cost estimates were obtained from Medicare claim files provided by the CMS. Drug costs were obtained from the 2005 Red Book. The resulting cost file produced by ABT Associates, Inc. provided an estimate of the costs of care associated with 9 Activities of Daily Living (ADL) categories, as shown in Table 23 (See Appendix 7).

For the purpose of the modeling exercise, two adjustments were made to the direct costs contained within the Slifka dataset owing mainly to the small number of respondents included in the sample. First, monitoring costs were estimated for patients with a confirmed diagnosis of MS but no physical or mental signs of impairment (ADL 1). These were assumed to include a general check-up, 1-2 visits to MS specialists, a blood count and an MRI scan. This monitoring cost was estimated to be approximately \$870 per patient per year. This monitoring cost was added to the costs of care for patients in ADL 1 estimated from the Slifka sample, resulting in a total estimated cost of \$3,196. Secondly, the costs associated with managing patients with very severe disease (ADL 9) appeared to be both underrepresented (n=1) and underestimated within the Slifka dataset. It is possible that this is a result of informative censoring, whereby those patients with more severe disease may be less likely to respond to the Slifka survey. In order to account for this potential bias, it was assumed that these patients would require indefinite nursing home care. The annual cost of a private room in a nursing home was taken from a survey reported by Metlife Market

Research.⁵⁶ Additional costs of MS-specific care for these patients were assumed to be similar to the non-hospitalization costs observed within the Slifka dataset. Indirect cost data were available from the Slifka dataset although these appeared to be underestimates; in the absence of more robust estimates within the literature, indirect costs were excluded from the cost-effectiveness analysis.

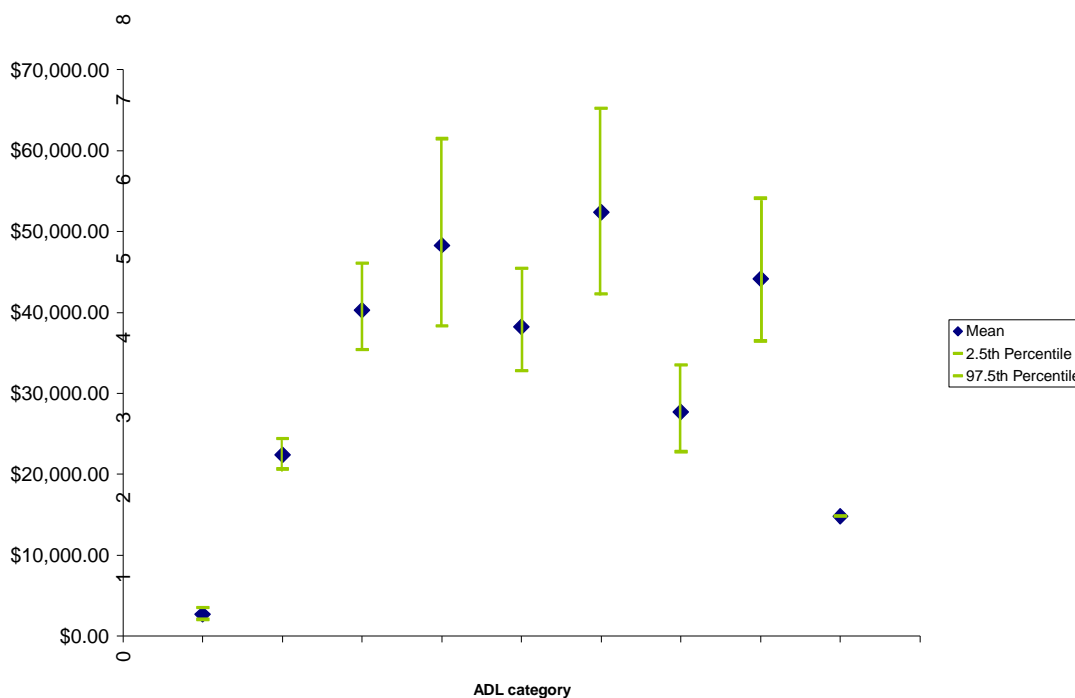
The mean costs associated with health care resources consumed according to ADL category from the Slifka dataset are shown in Table 23. These cost estimates include the assumptions for those patients who were minimally disabled and those patients who were severely disabled, as described above.

Table 23 Mean estimated annual cost of care according to ADL category within the Sonya Slifka dataset^{30;31}

ADL category	Unweighted number of respondents (2005 data)	Estimated direct costs
1-No MS Symptoms	3	\$3,126 [†]
2-Mild Symptoms, Non-Limiting	225	\$21,153
3-Mild Symptoms, Not Affecting Walking	298	\$35,791
4-Prob Walking, No Aid	107	\$44,743
5-25 Ft Without Aid	85	\$37,016
6-1 Side Support	111	\$51,641
7-2 Side Support	30	\$26,142
8-Wheelchair/Scooter	48	\$42,814
9-Bedridden	1	\$103,592 [†]
Table key		
[†] Derived through adjustments described above		

ABT Associates, Inc. kindly provided standard errors surrounding the mean number of times each resource item was utilized, as well as standard errors surrounding the proportion of patients utilizing each health care resource item. Uncertainty surrounding these cost estimates was explored by assigning lognormal distributions to describe the number of times each health care resource item was utilized, and beta distributions to describe the proportion of patients in the sample who utilized each resource item. Monte Carlo sampling methods (5,000 random samples of ADL-specific costs) were then used to estimate the 2.5th and 97.5th percentiles for the cost associated with each ADL category. Figure 5 shows the mean, 2.5th and 97.5% percentile costs associated with each ADL category.

Figure 5 Mean, 2.5th and 97.5th percentile health care resource costs associated with individual ADL categories



The Sonya Slifka dataset contains health care resource use associated with ADL categories, while the SchARR MS cost-effectiveness model describes disease severity in terms of the EDSS. Consequently, it was necessary to map the costs associated with individual ADL categories to the EDSS. As there were no EDSS-ADL paired data, the mapping had to be constructed descriptively rather than analytically. The EDSS is detailed in Appendix 1; a description of the ADL categories used within the Slifka dataset is presented in Appendix 7.

Mapping the Sonya Slifka data onto the EDSS

The first stage of the mapping process was to identify any ADL categories and EDSS states that were direct or partial matches. Two of the authors (CMcC and PT) undertook this assessment independently. Each assessor then identified a preferred match for each ADL state on the EDSS scale; the two assessments were then compared and a final agreed descriptive mapping was produced. As the EDSS has 20 states and ADL has only 9 states, it was necessary to map individual ADL categories to more than one EDSS state. The implication is that the cost analysis will be less sensitive than if EDSS-specific cost data were available.

Only two direct mappings for ADL categories and EDSS states were identified, although several other ADL categories were broadly reflected by EDSS states. ADL level 1 was mapped directly to EDSS 0, and ADL 9 was mapped directly to EDSS 9.0 and above. In addition, ADL 3 was descriptively similar to EDSS 3.0; both of these classifications describe a state of symptomatic MS that does not impact upon walking. ADL 8 is descriptively similar to EDSS 7.0, as both

classifications describe the reliance of the patient on a wheelchair. ADL 6 was mapped to EDSS 6.0, as unilateral support is a defining characteristic in both states. Similarly, ADL 7 was mapped to EDSS 6.5 as bilateral support is a defining characteristic in both states. Logically, ADL 2 lies between EDSS states 0 and 3.0. The cost-effectiveness model assumes that ADL 2 relates to the range of EDSS states between 1.0 and 2.5.

Mapping ADL categories 4 and 5 to specific EDSS states or ranges of states was problematic. The mapping decisions described above ensured that these two states lie somewhere between EDSS 3.5 and EDSS 5.5. However, ADL categories 4 and 5 are not logically distinct, as having problems walking without aid (ADL 4), and being able to walk 25 feet without aid (ADL 5) are distinct only to the degree that being able to walk more than 25 feet without aid is still considered to be a problem. This distinction is almost certainly dependent upon the nature of an individual's daily activities; if the individual never needs to walk more than 25 feet, then ADL 5 could be better than ADL 4. For this reason, cost data relating to ADL 4 and 5 were pooled. This pooled ADL category was mapped to the first EDSS state in which the patient is not being fully ambulatory; that is EDSS 5.0. The assumed mapping between the ADL categories and EDSS states is presented in Table 24.

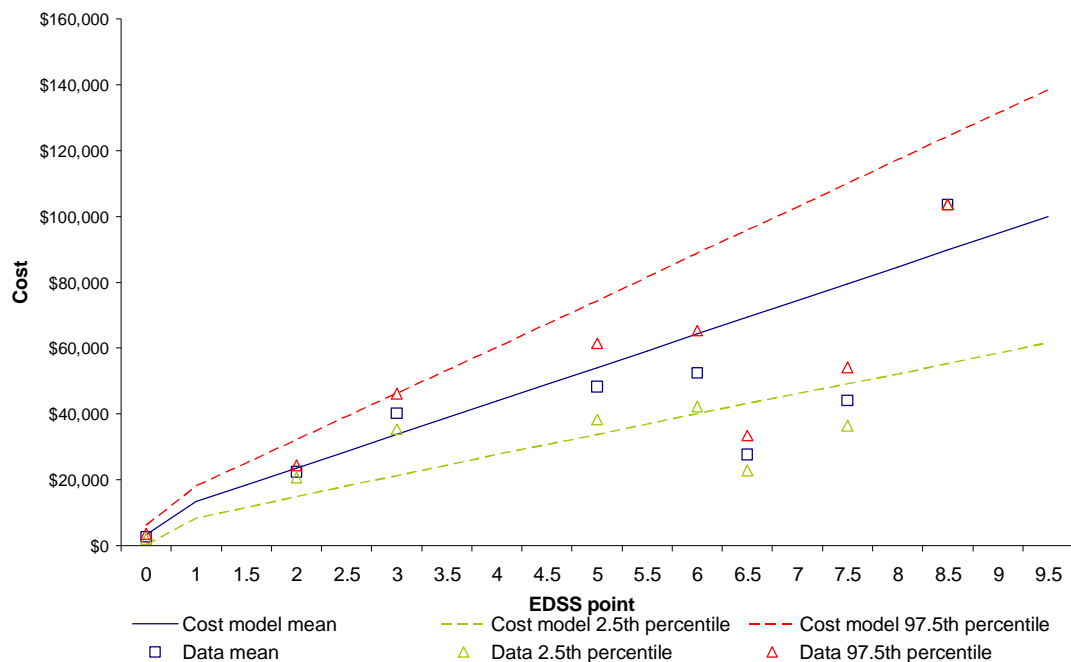
Table 24 Assumed mapping of ADL against EDSS

EDSS state ⁷	Sonya Slifka ADL category (2005 data)
0	1
1	2
1.5	2
2	2
2.5	2
3	3
3.5	3
4	3
4.5	3
5	Pooled 4 & 5
5.5	Pooled 4 & 5
6	6
6.5	7
7	8
7.5	8
8	8
8.5	8
9	9
9.5	9
10	Death due to MS (n/a)

As the Sonya Slifka data cover only a proportion of the range of EDSS health states, an EDSS-cost model was fitted to the Slifka data using simple linear regression analysis. In order to account

for potential bias in the small number of patients who populate the more severe ADL categories, greater weight was given to those ADL categories with a greater number of patients (ADL categories 1-5). Standard errors for the regression coefficients were estimated analytically using Monte Carlo sampling techniques in order to describe a similar degree of uncertainty surrounding the mean ADL cost estimates. Figure 6 shows the mean cost estimates from the Slifka dataset mapped onto the EDSS, together with the 2.5th and 97.5th percentile costs for each ADL category estimated from the probabilistic ADL cost model. These data are presented alongside the results of the regression model used to estimate the relationship between direct costs and increasing MS disability.

Figure 6 Mapped EDSS direct cost estimates and regression output including uncertainty estimates



Based upon the ADL-EDSS mapping process, the expected pattern of increasing resource use together with MS disability is not obvious from the Sonya Slifka dataset. This may be a result of the limitations in the relationship between the ADL and the EDSS, the exclusion of nursing care costs from the Slifka dataset, and the declining number of respondents in the more severe ADL categories. A simple straight line relationship was assumed between EDSS progression and the associated costs of care, and wide uncertainty was incorporated within the probabilistic sensitivity analysis to account for the uncertainty surrounding the relationship between costs of care and MS disability. The absence of robust evidence on the relationship between the EDSS and the costs of care in the US is clearly a limitation of this health economic analysis.

5.6.8.3 Costs associated with MS relapse

The Sonya Slifka dataset did not explicitly identify those health care resources consumed in the general costs of caring for individuals with MS in remission, and those costs associated with temporal relapses. However, as one of the principal measures of clinical efficacy of the disease-modifying therapies derives directly from their impact in terms of reducing the frequency of relapses (and consequently the costs associated with managing these exacerbations), it was important to delineate this cost from the general costs of MS care. The costs associated with the management of MS relapse may be incorporated within the cost estimates derived from the Slifka dataset, which may lead to the overestimation of EDSS-specific costs within the model.

The cost of managing a relapse was derived from a study reported by O'Brien and colleagues.⁵⁷ O'Brien et al undertook a direct cost analysis and cost modeling to estimate the costs associated with low-, medium- and high-level management of MS relapses based upon data obtained from all payer inpatient, ambulatory, and emergency room databases from several states, fee schedules, government reports, and associated literature.⁵⁷ O'Brien estimated that the cost of relapse varied considerably depending on the level of management required. The authors estimated that cost of managing a mild MS relapses was \$283, while the cost of medium- and high-level management of relapses were \$1,847 and \$12,870 respectively.⁵⁷ A weighted cost of relapse for use in the cost-effectiveness model was estimated by applying the proportion of mild, moderate and severe relapses observed within the placebo-controlled RCT of IFN β -1b 8MIU in the treatment of RRMS.¹⁵ This weighted cost of MS relapse was estimated to be \$3,158. This cost is similar to the cost of relapse assumed within the earlier 2001 assessment of disease-modifying therapies for MS undertaken on behalf of NICE.^{39;40}

5.6.9 Discounting

Good practice in health economic evaluation requires that costs and benefits that occur in the future are given less weight than those that occur in the present. Within the base case model, both costs and health outcomes are discounted at 5%.⁵⁸ Sensitivity analysis was also undertaken using discount rates of 0% and 3% for both costs and health effects.

5.6.10 Uncertainty analysis

The costs and effects of disease-modifying therapies in the management of RRMS and SPMS are subject to considerable uncertainty. The following sensitivity analyses were undertaken:

- § Structural scenario analysis to explore the impact of alternative assumptions concerning the appropriate time horizon for the analysis;
- § One-way sensitivity analysis to identify the key determinants of cost-effectiveness;
- § Scenario analysis using the mixed treatment comparisons which incorporated evidence concerning the comparative efficacy of IFN β -1b 8MIU, IFN β -1a 6MIU and IFN β -1a 44 μ g (See Sections 4.7-4.8, and Table 19).

- § Probabilistic (Monte Carlo) sensitivity analysis to explore the impact of second-order uncertainty surrounding mean parameter values on marginal cost-effectiveness outcomes.

5.6.10.1 Structural scenario analysis

The most appropriate time horizon for the economic evaluation of disease-modifying therapies in the management of MS is unclear; existing evaluations have used time horizons ranging from 2 years to the lifetime of MS patients. While the base case analysis assumes a lifetime horizon for the analysis, further scenario analysis was undertaken whereby time horizons of 5, 10, and 20 years were modeled. As the model assumes that benefits in terms of progression avoided are retained by patients who receive disease-modifying therapy even when they drop off therapy, longer time horizons present more favorable cost-effectiveness estimates, while shorter time horizons are less affected by this assumption, thus cost-effectiveness estimates are less favorable.

5.6.10.2 One-way sensitivity analysis

A number of one-way sensitivity analyses were undertaken to explore the impact of changing individual parameter values on central estimates of cost-effectiveness. First, the impact of treatment cessation upon progression to EDSS 7.0 on marginal cost-effectiveness estimates was considered. Due to the uncertainty surrounding the relationship between MS disability and the costs of care, a sensitivity analysis was undertaken whereby an exponential relationship between MS disability and costs of care was assumed, as reported for the earlier UK analysis.⁴¹ Illustrative sensitivity analysis was undertaken to explore the impact of avoiding relapse and progression on cost-effectiveness estimates. Other sensitivity analyses included varying the costs of relapse, varying drop-out rates, and varying assumptions concerning the discount rates for health effects and costs.

5.6.10.3 Scenario analysis using mixed treatment comparisons

The base case analysis presented within this report is restricted only to the placebo-controlled trials in RRMS and SPMS which met the inclusion criteria for this review. Further sensitivity analysis was undertaken to explore the impact of the modified treatment effectiveness estimates generated using mixed treatment comparisons (see Table 19) on cost-effectiveness estimates.

5.6.10.4 Probabilistic sensitivity analysis

In recent years, considerable research effort has been devoted to methods for describing uncertainty surrounding the costs and effects of healthcare interventions using probabilistic sensitivity analysis (See “Key Definitions”).⁵⁹⁻⁶² While confidence intervals are typically used to express uncertainty surrounding estimates of clinical efficacy, the use of confidence intervals to describe uncertainty surrounding mean cost-effectiveness ratios is problematic as the cost per QALY ratio does not exist on a continuous scale. More recently, the use of Cost Effectiveness

Acceptability Curves (CEACs) to represent parameter uncertainty has become standard practice in health economic evaluation. Each CEAC shows the likelihood that a treatment option produces the greatest amount of net benefit (See “Key Definitions”). A simple algorithm for generating CEACs is presented in Box 2.

Box 2 Algorithm for generating Cost Effectiveness Acceptability Curves

- (1) Describe second-order uncertainty surrounding all uncertain model parameters by assigning a unique parametric distribution (e.g. normal, lognormal, beta, Dirichlet);
- (2) Use Monte Carlo techniques to sample from all parameter distributions simultaneously over a large number of random iterations;
- (3) Calculate net benefits of each treatment option for a plausible range of willingness-to-pay values;
- (4) For each possible willingness-to-pay value, calculate the proportion of random samples for which the given treatment option results in the greatest amount of net benefit;
- (5) Repeat Step (4) for all treatment strategies.
- (6) Plot the probability of being cost-effective at each plausible willingness-to-pay threshold over the entire range of threshold values

Probabilistic sensitivity analysis was undertaken using the MS cost-effectiveness model in order to explore the impact of second-order uncertainty surrounding all mean parameter values on the marginal cost-effectiveness of each of the therapies. This was undertaken by describing parameter values within the model using probability distributions, and by randomly sampling from all uncertain distributions simultaneously using Monte Carlo simulation routines over 2,000 iterations. As noted throughout the systematic review of clinical effectiveness (See Chapter 4), the majority of published RCTs do not provide confidence intervals or standard errors surrounding relative relapse rates or relative hazard ratios for disease progression. In order to resolve this issue, the description of uncertainty surrounding treatment effectiveness outcomes was undertaken through a process of consultation with clinical experts, NICE and other stakeholders in 2001. Within this analysis, this description of uncertainty was based on outcome data reported within the published study reports and augmented using the earlier estimates derived within the 2001 NICE appraisal.³⁹

The results of the Monte Carlo sensitivity analysis are presented as marginal CEACs, which provide information on the likelihood that each individual disease-modifying therapy results in more net benefit than best supportive care alone. Probabilistic sensitivity analysis was undertaken for two of the analyses: (1) the base case model assumptions; and (2) the scenario in which all patients are assumed to discontinue treatment with disease-modifying therapy upon progression to EDSS 7.0.

supportive care; the resulting central estimate of marginal cost-effectiveness for IFN β -1a 6MIU versus best supportive care is estimated to be approximately \$112,500 per QALY gained. This estimate is slightly more favorable when IFN β -1a is self-administered, offering savings of around \$3,900 per patient over the equivalent treatment administered by a physician, which results in a marginal cost-effectiveness of approximately \$104,200 per QALY gained. However, it is crucial to note that the clinical efficacy evidence for IFN β -1a 6MIU available within the public domain was *not* based on an ITT analysis. IFN β -1a 22 μ g is expected to produce an additional 0.30 QALYs at an additional cost of \$59,800 when compared to best supportive care; this results in a mean cost-effectiveness estimate for IFN β -1a 22 μ g versus best supportive care of approximately \$198,500 per QALY gained. IFN β -1a 44 μ g is estimated to produce an additional 0.45 QALYs at an additional cost of \$59,000 when compared with best supportive care; the marginal cost-effectiveness for IFN β -1a 44 μ g versus best supportive care is estimated to be approximately \$131,900 per QALY gained.

Within its RRMS indication, the base case analysis suggests that IFN β -1b 8MIU is expected to produce an additional 0.32 QALYs at an additional cost of \$52,700 when compared to best supportive care; this results in a marginal cost-effectiveness ratio of \$164,100 per QALY gained. When disease-modifying therapy is assumed to continue upon progression to SPMS, IFN β -1b 8MIU is expected to produce an additional 0.49 QALYs at an additional cost of \$143,700 when compared to best supportive care alone; the estimated marginal cost-effectiveness ratio for IFN β -1b 8MIU versus best supportive care is estimated to be \$295,200 per QALY gained. GA 20mg is expected to produce an additional 0.15 QALYs at an additional cost of \$48,800; the resulting marginal cost-effectiveness ratio for GA 20mg versus best supportive care is estimated to be \$332,000.

The estimates presented in Table 25 were derived using the distributional mean values for all model parameters. As all model parameters are subject to ubiquitous uncertainty, Section 6.3 presents simple sensitivity analysis, including the modified treatment effectiveness estimates based upon the placebo-controlled and head-to-head trials of IFNB-1a and IFNB-1b presented in Chapter 4 (See Table 19). The impact of joint uncertainty surrounding the mean values of all model parameters on the marginal cost-effectiveness of these therapies was explored using probabilistic sensitivity analysis (See Section 6.4).

6.3 Simple sensitivity analysis

6.3.1 Impact of stopping treatment at EDSS 7.0

In clinical practice it is usual to continue treatment with disease-modifying therapy until the patient reaches a certain level of disability (around EDSS 6.5 or EDSS 7.0). However, owing to the absence of US clinical guidelines concerning this discontinuation strategy, the base case model

Table 31 shows that the alternative assumptions concerning discount rates for health outcomes and costs has only a minor impact upon the marginal cost per QALY gained for the RRMS treatment options.

6.3.5 Impact of dropout rates on cost-effectiveness estimates

Table 32 shows the impact of alternative assumptions concerning the therapy dropout rates.

Table 32 Impact of assumptions concerning drop-out rates on cost-effectiveness

Treatment option	Base case dropout rates (10% year 1, 10% year 2, 3% thereafter)	Dropout rates halved (5% year 1, 5% year 2, 1.5% thereafter)	Dropout rates doubled (20% year 1, 20% year 2, 6% thereafter)
Physician-administered IFN β -1a 6MIU (Avonex)	\$112,531	\$114,188	\$109,640
Self-administered IFN β -1a 6MIU (Avonex)	\$104,199	\$105,665	\$101,626
IFN β -1a 22 μ g (Rebif)	\$198,483	\$202,324	\$191,924
IFN β -1a 44 μ g (Rebif)	\$131,949	\$134,078	\$128,266
IFN β -1b 8MIU (Betaseron) for RRMS	\$164,096	\$167,160	\$158,841
GA 20mg (Copaxone)	\$332,006	\$337,710	\$322,369
IFN β -1b 8MIU (Betaseron) for RRMS and SPMS	\$295,186	\$316,972	\$261,081
Best supportive care	-	-	-

Table 32 shows that the assumptions used within the base case analysis have only a minor impact upon the estimates of marginal cost-effectiveness for the disease-modifying therapies for RRMS. Analyses in which lower dropout rates are assumed would be expected to result in less favorable marginal cost-effectiveness estimates.

6.3.7 Impact of relative relapse rates and relative hazard ratios for disease progression on cost-effectiveness estimates

Tables 33 and 34 demonstrate the impact of alternative assumptions concerning disease progression and relapse rates on the estimates of marginal cost-effectiveness for the disease-modifying therapies compared to best supportive care. As confidence intervals and standard errors surrounding the relative hazard ratios for disease progression and relative relapse rates were not available from the trial publications (See Section 4.7), the percentage improvements in hazard ratios presented here are arbitrary; the intention of this sensitivity analysis is to illustrate the impact of potential improvements in disease progression on the marginal cost-effectiveness of these therapies compared to best supportive care.

Table 33 Illustrative sensitivity analysis surrounding the impact of reduction in disease progression on marginal cost-effectiveness estimates

Treatment option	Percentage reduction in relative hazard ratio for disease progression [†]			
	Base case	10% reduction	25% reduction	50% reduction
Physician-administered IFN β -1a 6MIU (Avonex)	\$112,531	\$95,745	\$76,287	\$53,049
Self-administered IFN β -1a 6MIU (Avonex)	\$104,199	\$88,501	\$70,259	\$48,383
IFN β -1a 22 μ g (Rebif)	\$198,483	\$153,993	\$111,928	\$70,830
IFN β -1a 44 μ g (Rebif)	\$131,949	\$111,249	\$87,886	\$60,805
IFN β -1b 8MIU (Betaseron) for RRMS	\$164,096	\$129,082	\$94,753	\$59,969
GA 20mg (Copaxone)	\$332,006	\$204,514	\$125,384	\$68,248
IFN β -1b 8MIU (Betaseron) for RRMS and SPMS	\$295,186	\$244,745	\$192,015	\$131,310
Best supportive care	-	-	-	-

[†] Note: The base case progression hazard for IFN β -1a 6MIU was estimated to be 0.58, therefore a 10% improvement in this ratio is estimated to be 0.52.

Table 34 Illustrative sensitivity analysis surrounding the impact of reduction in relapse rates on marginal cost-effectiveness estimates

Treatment option	Percentage reduction in relative relapse rate			
	Base case	10% reduction	25% reduction	50% reduction
Physician-administered IFN β -1a 6MIU (Avonex)	\$112,531	\$109,727	\$105,649	\$99,173
Self-administered IFN β -1a 6MIU (Avonex)	\$104,199	\$101,498	\$97,570	\$91,332
IFN β -1a 22 μ g (Rebif)	\$198,483	\$193,386	\$186,046	\$174,561
IFN β -1a 44 μ g (Rebif)	\$131,949	\$129,299	\$125,429	\$119,246
IFN β -1b 8MIU (Betaseron) for RRMS	\$164,096	\$159,909	\$153,859	\$144,350
GA 20mg (Copaxone)	\$332,006	\$317,910	\$298,362	\$269,439
IFN β -1b 8MIU (Betaseron) for RRMS and SPMS	\$295,186	\$288,876	\$279,769	\$265,472
Best supportive care	-	-	-	-

The hypothetical results presented in Tables 33 and 34 demonstrate that reductions in disease progression have a substantial impact upon the cost-effectiveness of these therapies, while improvements in relapse rates do not have a marked impact. For example, reducing the relative hazard ratio for disease progression for IFN β -1b 8MIU by 50% would reduce the mean cost-effectiveness estimate from \$164,100 to \$60,000 per QALY gained. However, reducing the relative relapse rate for IFN β -1b 8MIU by 50% would reduce the mean cost-effectiveness estimate to \$144,400 per QALY gained.

6.3.7 Impact of relapse cost on cost-effectiveness estimates

Table 35 demonstrates the impact of alternative assumptions concerning the cost of managing relapses on the central estimates of cost-effectiveness, as compared to the base case assumption whereby the mean cost of managing a relapse episode was estimated to be approximately \$3,158.

Table 35 Impact of cost of relapse on central estimates of cost-effectiveness

Treatment option	Base case (relapse cost = \$3,158)	Doubled relapse cost (relapse cost = \$6,316)	Halved relapse cost (relapse cost = \$1,579)
Physician-administered IFN β -1a 6MIU (Avonex)	\$112,531	\$111,827	\$112,883
Self-administered IFN β -1a 6MIU (Avonex)	\$104,199	\$103,494	\$104,551
IFN β -1a 22 μ g (Rebif)	\$198,483	\$193,292	\$201,079
IFN β -1a 44 μ g (Rebif)	\$131,949	\$128,441	\$133,703
IFN β -1b 8MIU (Betaseron) for RRMS	\$164,096	\$158,982	\$166,653
GA 20mg (Copaxone)	\$332,006	\$318,748	\$338,634
IFN β -1b 8MIU (Betaseron) for RRMS and SPMS	\$295,186	\$290,253	\$297,652
Best supportive care	-	-	-

This sensitivity analysis clearly demonstrates that the cost associated with managing MS relapses is not an important determinant of the long-term cost-effectiveness of the disease-modifying therapies.

6.3.6 Impact of alternative EDSS-cost relationship on cost-effectiveness estimates

As noted in Section 5.6.8.2, there is considerable uncertainty surrounding the relationship between the EDSS and the costs of care. The base case analysis assumes a straight-line relationship between the costs of care and increasing disability as measured by the EDSS. The UK cost-effectiveness model did not employ this assumption, but instead assumed that costs increase exponentially alongside worsening on the EDSS. A further sensitivity analysis was undertaken to explore the impact of making this assumption on the estimates of marginal cost-effectiveness IFN β and GA in the US. The costs of care in the UK for each EDSS state were estimated relative to the cost of EDSS 9.5. These EDSS-specific multipliers were then multiplied by the cost of EDSS 9.5 assumed within the US analysis in order to estimate the cost of care associated with each EDSS state. This alternative exponential EDSS-cost function is shown in Table 36.

The results presented in Tables 33, 34, 36 and 37 clearly suggest that the impact of the disease-modifying therapies on disease progression is a key determinant of their cost-effectiveness, while reductions in the incidence of relapse do not substantially affect cost-effectiveness outcomes. This does not mean that the management of relapse is not clinically important, but rather that the effect of the disease-modifying therapies in terms of reducing relapse rates does not have a marked impact upon their marginal effects and costs. This is an important finding; although the disease-modifying therapies have all demonstrated a significant reduction in relapse rates compared to placebo (See Section 4.8), such treatment benefits offer only minor cost savings and health improvements within the broader management of the disease. The analysis of resource use and cost data suggests that the costs of disease management generally increase as EDSS worsens, while the EDSS utility data suggests that HRQoL decreases as EDSS increases. The results presented above are intuitively sensible, as the effect of slowing disease progression delays the progression to more disabling states of health which are more expensive to manage and which have a more detrimental impact upon the quality of life of patients. Crucially, while this analysis demonstrates that the impact of disease-modifying therapy on disease progression has a considerable influence upon resulting cost-effectiveness estimates, none of these therapies have demonstrated significant improvements in delaying time to sustained disease progression within a robust, long-term trial setting.

6.4 Probabilistic sensitivity analysis

6.4.1 Probabilistic sensitivity analysis using the base case assumptions

Table 38 shows the degree of uncertainty surrounding the mean costs and effects of each of the disease modifying therapies as well as best supportive care. It should be noted that the standard deviations reflect not only the uncertainty surrounding treatment effects, but rather the uncertainty surrounding *all* of the model parameters.

Table 38 Uncertainty surrounding mean costs and health outcomes for disease-modifying therapies and best supportive care – base case scenario

Treatment option	Cost (SD)	QALYs (SD)
Physician-administered IFN β -1a 6MIU (Avonex)	\$806,762 (\$142,420)	2.87 (1.10)
Self-administered IFN β -1a 6MIU (Avonex)	\$802,823 (\$142,308)	2.87 (1.11)
IFN β -1a 22 μ g (Rebif)	\$813,328 (\$143,038)	2.70 (1.09)
IFN β -1a 44 μ g (Rebif)	\$812,547 (\$142,732)	2.85 (1.10)
IFN β -1b 8MIU (Betaseron) for RRMS	\$806,241 (\$142,899)	2.72 (1.09)
GA 20mg (Copaxone)	\$802,334 (\$143,123)	2.55 (1.07)
IFN β -1b 8MIU (Betaseron) for RRMS and SPMS	\$897,264 (\$159,565)	2.89 (1.17)
Best supportive care	\$753,564 (\$142,802)	2.40 (1.04)

Table 38 clearly suggests that the costs and effects of each of the disease-modifying therapies are subject to a considerable degree of uncertainty. Standard deviations for mean costs range from around \$142,300 to \$159,600, while standard deviations surrounding mean QALY gains range

Figure 8 Marginal CEACs for disease-modifying therapies compared to best supportive care alone – treatment discontinuation at EDSS 7.0

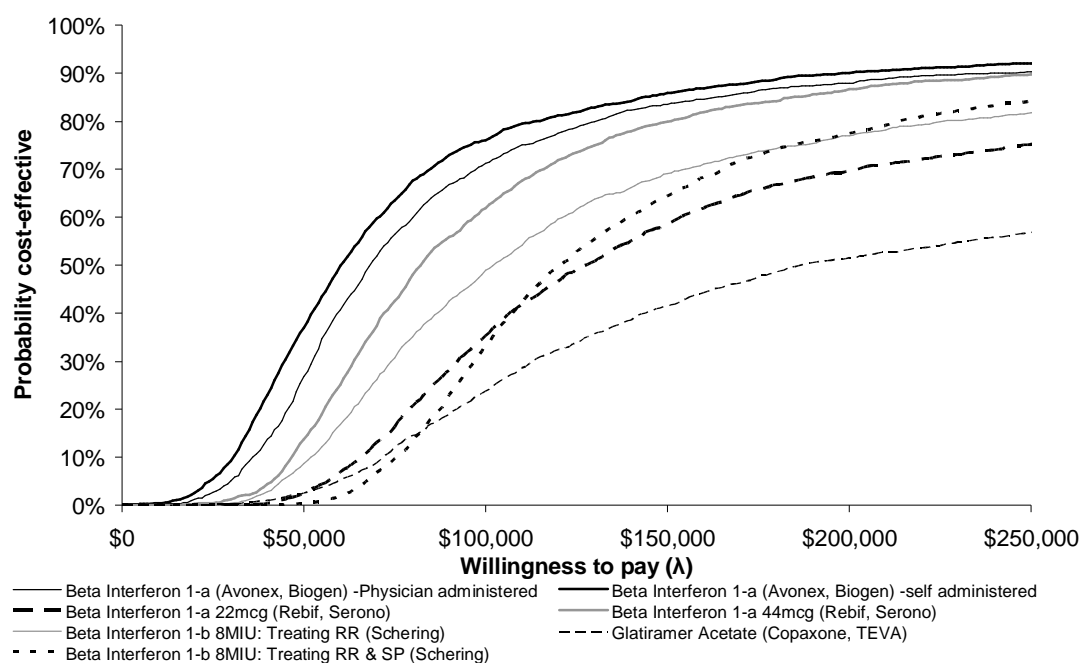


Figure 8 suggests that when patients discontinue treatment upon progression to EDSS 7.0, the probability that any of the disease-modifying therapies has a marginal cost-effectiveness that is better than \$60,000 per QALY gained when compared to best supportive care alone is 0.50 or lower. The probability that any of the disease-modifying therapies has a marginal cost-effectiveness that is better than \$100,000 per QALY gained when compared to best supportive care alone is around 0.76 or lower. To further illustrate the results of the CEACs presented in Figure 8, Table 41 shows the probability that each of the disease-modifying therapies has a cost-effectiveness that is better than \$60,000 per QALY gained and \$100,000 per QALY gained respectively when compared against best supportive care, assuming that patients drop off therapy upon progression to EDSS 7.0.

Table 41 Tabular description of marginal CEACs – treatment discontinuation at EDSS 7.0

Treatment option	Probability cost-effectiveness is better than \$60,000 per QALY gained	Probability cost-effectiveness is better than \$100,000 per QALY gained
Physician-administered IFNβ-1a 6MIU (Avonex)	40.67%	71.24%
Self-administered IFNβ-1a 6MIU (Avonex)	49.87%	76.09%
IFNβ-1a 22μg (Rebif)	6.90%	35.32%
IFNβ-1a 44μg (Rebif)	25.16%	61.98%
IFNβ-1b 8MIU (Betaseron) for RRMS	16.86%	48.97%
GA 20mg (Copaxone)	5.35%	23.81%
IFNβ-1b 8MIU (Betaseron) for RRMS and SPMS	1.90%	33.37%
Best supportive care	-	-

disease progression and relapse would be useful. The long-term monitoring of the experience of individuals with MS who are receiving disease-modifying therapies is currently being monitored as part of the UK Department of Health's MS risk-sharing and monitoring scheme. It is anticipated that this study will follow-up patients receiving these disease-modifying therapies for up to 10 years.

- There is a dearth of evidence concerning the effectiveness of sequences of these therapies. Further research concerning the clinical effectiveness and cost-effectiveness of alternative sequences of disease-modifying therapies is merited.
- The Sonya Slifka dataset is clearly a rich and valuable source of health care utilization data for individuals with MS. However, the transposition of these resource use data which are based upon a specific ADL scale, onto the EDSS is problematic. The absence of robust cost evidence to inform cost-effectiveness models of the IFN β has been noted within previous cost-effectiveness models developed in the US.³⁷ Further research on the relationship between the EDSS, health care utilization and costs of MS care would be valuable.
- While the Canadian HUI3 data are highly consistent with other utility sources from other countries, it is possible that US valuations may differ from those used in this analysis. Further information concerning the relationship between EDSS and health utilities within the US population may be valuable.

Appendix 2 Worked example of progression-free survival hazards

The calculation of relative EDSS progression hazards assumes an exponential survivor function whereby

$$S(t) = e^{-\alpha t} \quad [1]$$

where $S(t)$ is the probability of survival at a given time-point, α is the hazard of EDSS progression, and t is time.

This worked example shows the calculation of the relative EDSS progression hazard for IFN β -1a 6MIU (Avonex) versus placebo.

The text reported in publication of IFN β -1a 6MIU (Avonex) versus placebo states that *“The proportion with progression of disability by 104 weeks estimated from Kaplan-Meier curves was 34.9% in placebo recipients and 21.9% in interferon beta-1a recipients.”*

Therefore at 2 years, the proportion of patients who were progression-free in the intervention arm was 78.1% or $S(t) = 0.781$, while the proportion of patients who were progression-free in the placebo arm was 65.1% or $S(t) = 0.651$.

By rearranging the survivor function for $S(t)$ shown in equation [1], we find that

$$\alpha = -\frac{1}{t} \ln \frac{1}{S(t)} \quad [2]$$

Therefore at 2 years, the hazard for IFN β -1a 6MIU (α_1) = $-(1/2) * \ln(1/0.781) = -0.12359$

If we do the same calculation for the placebo arm, the hazard for EDSS progression in this arm (α_2) is estimated to be $-(1/2) * \ln(1/0.651) = -0.21462$.

As the relative hazard is calculated to be α_1/α_2 , the relative progression hazard for IFN β -1a 6MIU (Avonex) versus placebo is $-0.12359/-0.21462 = 0.5758$.

Appendix 4 List of included and excluded studies

List of included trials

Trial	Intervention	Intervention 2	Placebo	Type of MS	N
Comi ¹⁴	GA 20mg Copaxone		placebo injection	RR	239
Johnson ²⁰	GA 20mg Copaxone		placebo injection	RR	251
PRISMS ²³	IFNB-1a 22µg Rebif	IFNB-1a 44µg Rebif	placebo injection	RR	560
INCOMIN ²⁷	IFNB-1b 8MIU (0.25mg) Betaseron	IFNB-1a 6MIU/30µg Avonex	none	RR	188
EVIDENCE ²⁸	IFNB-1a 44µg Rebif	IFNB-1a 6MIU/30µg Avonex	none	RR	677
Jacobs ²¹ / Rudick ⁶⁴	IFNB-1a 6MIU/30µg Avonex		placebo injection	RR	301
IFNBMSSG ¹⁵	IFNB-1b 8MIU (0.25mg) Betaseron	(IFNB-1b 1.6mIU)	placebo injection	RR	372
European Study Group ²⁴	IFNB-1b 8MIU (0.25mg) Betaseron		placebo injection	SP	718
NASPMS ²⁶	IFNB-1b 8MIU (0.25mg) Betaseron	(IFNB-1b 5mIU per m squared, body surface area)	placebo injection	SP	939

Appendix 5 List of model parameters

Parameter group	Parameter name	Distribution	Parameter 1	Parameter 2	Parameter 3
All Cause mortality	Probability female	Normal	0.67	0.01	
Initial State Distribution	StartRR0	Normal	221.00		
	StartRR1-0	Normal	55.25		
	StartRR1-5	Normal	55.25		
	StartRR2-0	Normal	55.25		
	StartRR2-5	Normal	55.25		
	StartRR3-0	Normal	143.25		
	StartRR3-5	Normal	143.25		
	StartRR4-0	Normal	143.25		
	StartRR4-5	Normal	143.25		
	StartRR5-0	Normal	215.00		
	StartRR5-5	Normal	215.00		
	StartRR6-0	Normal	308.00		
	StartRR6-5	Normal	123.00		
	StartRR7-0	Normal	67.50		
	StartRR7-5	Normal	67.50		
	StartRR8-0	Normal	67.50		
	StartRR8-5	Normal	67.50		
StartRR9-0	Normal	3.50			
StartRR9-5	Normal	3.50			
StartRR10-0	Normal	0.00			
EDSS cost function	Gradient	Normal	\$10,169	\$2,000	
	Intercept	Normal	\$3,126	\$1,500	
Relapse cost	Cost of relapse	Lognormal	7.56	1.00	
DMT costs	Cost Avonex	N/a	\$12,438.57		
	Cost Rebif 22mcg	N/a	\$13,965.83		
	Cost Rebif 44mcg	N/a	\$13,965.83		
	Cost Betaseron	N/a	\$12,344.03		
	Cost Copaxone	N/a	\$11,613.38		
Utility curve parameters	A	Beta	32.13	11.87	-0.59
	alpha	Lognormal	-0.54	0.19	
	beta	Lognormal	-0.99	0.22	
Relapse utility	Disutility for relapse	Normal	0.22	0.09	
Relapse duration	Relapse duration (days)	Normal	46.00	10.00	
Beta - sojourn times	RR0	Lognormal	5.50	0.34	
	RR1	Lognormal	6.20	0.47	
	RR1.5	Lognormal	6.20	0.47	
	RR2	Lognormal	3.49	0.11	
	RR2.5	Lognormal	3.49	0.11	
	RR3	Lognormal	2.24	0.09	
	RR3.5	Lognormal	2.24	0.09	
	RR4	Lognormal	1.56	0.14	
	RR4.5	Lognormal	1.56	0.14	
	RR5	Lognormal	1.33	0.13	
	RR5.5	Lognormal	1.33	0.13	
RR6	Lognormal	1.61	0.14		

	RR6.5	Lognormal	1.61	0.14	
	RR7	Lognormal	1.40	0.23	
	RR7.5	Lognormal	1.40	0.23	
	RR8	Lognormal	2.47	0.43	
	RR8.5	Lognormal	2.47	0.43	
	RR9	Lognormal	1.45	0.83	
	RR9.5	Lognormal	1.45	0.83	
Relative hazards of EDSS Progression	Relative risk Avonex	Lognormal	0.58	0.19	
	Relative risk 22mcg Rebif	Lognormal	0.72	0.19	
	Relative risk 44mcg Rebif	Lognormal	0.60	0.19	
	Relative risk Betaseron RRMS	Lognormal	0.71	0.18	
	Relative risk Copaxone	Lognormal	0.86	0.23	
	Relative risk Betaseron SPMS	Lognormal	0.72	0.18	
Relative risk of relapse	Relative risk Avonex	Lognormal	0.82	0.125	
	Relative risk 22mcg Rebif	Lognormal	0.71	0.084	
	Relative risk 44mcg Rebif	Lognormal	0.68	0.084	
	Relative risk Betaseron RRMS	Lognormal	0.70	0.09	
	Relative risk Copaxone	Lognormal	0.70	0.105	
	Relative risk Betaseron SPMS	Lognormal	0.69	0.09	
Adverse event parameters	Proportion of patients experiencing AEs	Beta	1.28	2.70	
	AE utility decrement	Beta	0.15	2.80	
Dropouts	Year 1,2 dropouts	Beta	0.80	7.20	
	Subsequent dropouts	Beta	0.30	10.00	
Relapse count	RelRR0	Lognormal	-0.12	0.30	
	RelRR1	Lognormal	-0.24	0.30	
	RelRR1.5	Lognormal	-0.38	0.30	
	RelRR2	Lognormal	-0.43	0.30	
	RelRR2.5	Lognormal	-0.40	0.30	
	RelRR3	Lognormal	-0.49	0.30	
	RelRR3.5	Lognormal	-0.50	0.30	
	RelRR4	Lognormal	-0.59	0.30	
	RelRR4.5	Lognormal	-0.64	0.30	
	RelRR5	Lognormal	-0.64	0.30	
	RelRR5.5	Lognormal	-0.63	0.30	
	RelRR6	Lognormal	-0.66	0.30	
	RelRR6.5	Lognormal	-0.65	0.30	
	RelRR7	Lognormal	-0.80	0.30	

	RelRR7.5	Lognormal	-0.86	0.30	
	RelRR8	Lognormal	-1.00	0.30	
	RelRR8.5	Lognormal	-1.13	0.30	
	RelRR9	Lognormal	-1.22	0.30	
	RelRR9.5	Lognormal	-1.30	0.30	
	RelSP2	Lognormal	-0.50	0.30	
	RelSP2.5	Lognormal	-0.62	0.30	
	RelSP3	Lognormal	-0.66	0.30	
	RelSP3.5	Lognormal	-0.72	0.30	
	RelSP4	Lognormal	-0.72	0.30	
	RelSP4.5	Lognormal	-0.79	0.30	
	RelSP5	Lognormal	-0.86	0.30	
	RelSP5.5	Lognormal	-0.92	0.30	
	RelSP6	Lognormal	-1.02	0.30	
	RelSP6.5	Lognormal	-1.11	0.30	
	RelSP7	Lognormal	-1.20	0.30	
	RelSP7.5	Lognormal	-1.28	0.30	
	RelSP8	Lognormal	-1.38	0.30	
	RelSP8.5	Lognormal	-1.45	0.30	
	RelSP9	Lognormal	-1.53	0.30	
	RelSP9.5	Lognormal	-1.59	0.30	
Mean sojourn times in DSS states	RR0	Lognormal	1.11	0.18	
	RR1	Lognormal	1.81	0.09	
	RR2	Lognormal	1.88	0.05	
	RR3	Lognormal	1.30	0.06	
	RR4	Lognormal	0.79	0.09	
	RR5	Lognormal	0.72	0.10	
	RR6	Lognormal	1.26	0.11	
	RR7	Lognormal	1.05	0.19	
	RR8	Lognormal	2.19	0.34	
	RR9	Lognormal	1.45	0.83	
	SP2	Lognormal	0.98	0.16	
	SP3	Lognormal	0.95	0.09	
	SP4	Lognormal	0.52	0.09	
	SP5	Lognormal	1.05	0.08	
	SP6	Lognormal	1.42	0.07	
	SP7	Lognormal	1.68	0.08	
	SP8	Lognormal	1.41	0.11	
	SP9	Lognormal	1.40	0.21	
	DSS - EDSS ratios	RR1	Lognormal	-0.23	0.26
RR2		Lognormal	0.34	0.26	
RR3		Lognormal	-0.04	0.26	
RR4		Lognormal	0.35	0.40	
RR5		Lognormal	0.18	0.40	
RR6		Lognormal	0.00	0.40	
RR7		Lognormal	0.00	0.40	
RR8		Lognormal	0.00	0.40	
RR9		Lognormal	0.00	0.40	
SP2		Lognormal	0.00	0.40	
SP3		Lognormal	0.17	0.26	
SP4		Lognormal	0.16	0.20	
SP5		Lognormal	0.03	0.20	
SP6		Lognormal	-0.06	0.20	
SP7		Lognormal	0.00	0.40	
SP8	Lognormal	0.00	0.40		
SP9	Lognormal	0.00	0.40		

Table key		Distribution	Parameter 1	Parameter 2	Parameter 3
		Normal	Mean	Standard error	N/a
		Lognormal	Log of mean	Log standard error	“Smoothing” coefficient
		Beta	Number of successes	Number of trials	“Smoothing” coefficient

Appendix 6 Disability status scale

DSS	Description
1	No disability & minimal neurologic sign
2	Minimal disability - slight weakness or stiffness, mild disturbance of gait or mild visual disturbance
3	Moderate disability - monoparesis (partial or incomplete paralysis affecting one or part of one extremity) mild hemiparesis (slight paralysis affecting one side of body) moderate ataxia, disturbing sensory loss, prominent urinary or eye symptom, or a combination of lesser dysfunction
4	Relatively severe disability, but fully ambulatory without aid, self sufficient and able to be up and about 12 hours a day, does not prevent the ability to work or carry on normal living activities, excluding sexual dysfunction
5	Disability is severe enough to preclude working, maximal motor function involves walking unaided up to 500 meters
6	Needs assistance walking, for example a cane, crutches, or braces
7	Essentially restricted to a wheelchair but able to wheel oneself and enter and leave the chair without assistance
8	Essentially restricted to bed or a chair, retains many self care functions and has effective use of arms
9	Helpless and bedridden
10	Death due to MS - results from respiratory paralysis, coma of uncertain origin, or following repeated or prolonged epileptic seizures

Source: Kurtzke JF. On the evaluation of disability in multiple sclerosis. *Neurology*. 1961; 11(8).

Appendix 7 ADL categories used within the Sonya Slifka dataset

ADL category	Description
1	No MS Symptoms
2	Mild Symptoms, Non-Limiting
3	Mild Symptoms, Not Affecting Walking
4	Problems Walking, No Aid
5	25 Ft Without Aid
6	1 Side Support
7	2 Side Support
8	Wheelchair/Scooter
9	Bedridden

References

1. Compston A. *McAlpine's multiple sclerosis*. London: Churchill Livingstone, 1998.
2. The National Multiple Sclerosis Society. About MS. 2005. Available from <http://www.nmss.org/> [Accessed 03/07/2005].
3. National Institute for Clinical Excellence. Multiple Sclerosis - National clinical guidelines for NHS management in primary and secondary care. 2003. London, NICE.
4. Richards RG, Sampson FC, Beard SM, Tappenden P. A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models. *Health Technology Assessment* 2002;6:i-73.
5. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD *et al*. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of Neurology* 2001;50:121-7.
6. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L *et al*. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria." *Annals of Neurology* 2005;58:840-6.
7. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale. *Neurology* 1983;33:1444-52.
8. Whetten-Goldstein K, Sloan FA, Goldstein LB, Kulas ED. A comprehensive assessment of the cost of multiple sclerosis in the United States. *Multiple Sclerosis* 1998;4:419-25.
9. Beard SM, Hunn A, Wight JP. Treatments for pain and spasticity in multiple sclerosis: A rapid and systematic review. *Health Technology Assessment* 2003;7:i-124.
10. Centre for Drug Evaluation and Research. Catalog of FDA approved drug products. 2005 US Food and Drug Administration. Available from: <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/> [Accessed 08/03/2005].
11. Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness: Study quality assessment. 2001. York, CRD.
12. Ford CC, Johnson KP, Lisak RP, Panitch HS, Shifroni G, Wolinsky JS *et al*. A prospective open-label study of glatiramer acetate: over a decade of continuous use in multiple sclerosis patients. *Multiple Sclerosis* 2006;12:309-20.

13. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
14. Comi G, Filippi M, Wolinsky JS and the 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. *Annals of Neurology* 2001;49:290-7.
15. IFNB Multiple Sclerosis Study Group. Interferon beta-1b in the treatment of multiple sclerosis: Final outcome of the randomized controlled trial. *Neurology* 1995;45:1277-85.
16. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov chain Monte Carlo in practice*. London: Chapman and Hall, 1996.
17. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 2004;23:3105-24.
18. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet* 1999;354:1896-900.
19. Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP *et al*. Copolymer 1 reduces relapse rate and improves disability in relapsing- remitting multiple sclerosis: Results of a phase III multicenter, double- blind, placebo-controlled trial. *Neurology* 1995;45:1268-76.
20. Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP and the 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;50:701-8.
21. Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM and the Multiple Sclerosis Collaborative Research Group. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Annals of Neurology* 1996;39:285-94.
22. 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;43:655-61.

23. PRISMS Study Group. Randomised double-blind placebo-controlled study of interferon beta -1a in relapsing/remitting multiple sclerosis. *Lancet* 1998;352:1498-504.
24. European Study Group on Interferon beta in secondary progressive MS. Placebo-controlled multicentre randomised trial of interferon beta-1b in treatment of secondary progressive multiple sclerosis. *Lancet* 1998;352:1491-7.
25. Kappos L, Polman C, Pozzilli C, Thompson A, Beckmann K, Dahlke F *et al.* Final analysis of the European multicenter trial on IFNbeta-1b in secondary-progressive MS. *Neurology* 2001;57:1969-75.
26. FDA and Unger, EF. Medical Officer Clinical Review STN 103471. FDA. 2005.
27. Durelli L, Verdun E, Barbero P, Bergui M, Versino E, Ghezzi A *et al.* Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: results of a 2-year prospective randomised multicentre study (INCOMIN). *Lancet* 2002;359:1453-60.
28. Panitch H, Goodin DS, Francis G, Chang P, Coyle PK, O'Connor P *et al.* Randomized, comparative study of interferon beta-1a treatment regimens in MS: The EVIDENCE Trial. *Neurology* 2002;59:1496-506.
29. Freeman JA, Thompson AJ, Fitzpatrick R, Hutchinson M, Miltenburger C, Beckmann K *et al.* Interferon-beta 1b in the treatment of secondary progressive MS: impact on quality of life. *Neurology* 2001;57:1870-5.
30. Minden SL, Frankel D, Hadden LS, Srinath KP, Perloff JN. Disability in elderly people with multiple sclerosis: An analysis of baseline data from the Sonya Slifka Longitudinal Multiple Sclerosis Study. *Neurorehabilitation* 2004;19:55-67.
31. Minden SL, Frankel D, Hadden L, Perloff J, Srinath KP, Hoaglin D. The Sonya Slifka Longitudinal Multiple Sclerosis Study: Methods and sample characteristics. *Multiple Sclerosis* 2006; 12:24-38.
32. Brown MG. HUI3 utility data from cohort of patients in Nova Scotia. Data held on file 2005.
33. Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D. A cost-utility analysis of beta interferon for multiple sclerosis. *Health Technology Assessment* 1998;2:i-66.
34. Brown MG, Murray TJ, Sketris IS, Fisk JD, LeBlanc JC, Schwartz CE *et al.* Cost-effectiveness of interferon beta-1b in slowing multiple sclerosis disability progression. First estimates. *International Journal of Technology Assessment in Health Care* 2000;16:751-67.

35. Forbes RB, Lees A, Waugh N, Swingler RJ. Population based cost utility study of interferon beta-1b in secondary progressive multiple sclerosis. *BMJ* 1999;319:1529-33.
36. Kendrick M, Johnson KP. Long term treatment of multiple sclerosis with interferon beta may be cost effective. *Pharmacoeconomics* 2001;18:45-53.
37. Prosser, L. The cost effectiveness of treatments for multiple sclerosis. 2000. Cambridge, MA, Harvard University.
38. National Institute for Clinical Excellence. Review of independent and sponsored models on the cost effectiveness of disease modifying therapies in Multiple Sclerosis: Presentation to stakeholders on cost-effectiveness models for beta interferon and glatiramer acetate. Minutes of the Technology Appraisals Committee meeting. 13th December 2000. Available from: <http://www.nice.nhs.uk/page.aspx?o=16091> [Accessed 04/08/2005].
39. Tappenden P, Chilcott JB, O'Hagan, A, McCabe C, Cooper NJ, Abrams KA, *et al.* Cost-effectiveness of beta interferon and glatiramer acetate in the management of multiple sclerosis. Report to the National Institute for Clinical Excellence. 2001.
40. Chilcott JB, McCabe C, Tappenden P, O'Hagan A, Cooper NJ, Abrams KA *et al.* Modelling the cost effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis. *BMJ* 2003;326.
41. Tappenden P, Chilcott JB, Eggington S, Oakley JE, McCabe C. Methods for expected value of information analysis for computationally expensive models. Developments on the health economics of beta interferon and glatiramer acetate for multiple sclerosis. *Health Technology Assessment* 2004;8:i-92.
42. Ebers, G. London Ontario cohort study: Confidential data on file at National Institute for Clinical Excellence. London. 2001.
43. Sonnenberg FA, Beck RB. Markov models in medical decision making: A practical guide. *Medical Decision Making* 1993;13:332-8.
44. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998;13:397-409.
45. Association of British Neurologists. Guidelines for treatment with interferon beta and glatiramer acetate in multiple sclerosis. 2001. London, ABN.
46. Kurtzke JF. On the evaluation of disability in multiple sclerosis. *Multiple Sclerosis* 1961;11.

47. Weinschenker B, Bass B, Rice G, Noseworthy J, Carriere W, Baskerville J. The natural history of multiple sclerosis: a geographically based study. I Clinical course and disability. *Brain* 1989;112:133-46.
48. Goodkin D, Hertsgaard D, Rudick RA. Exacerbation rates and adherence to disease type in a prospectively followed-up population with multiple sclerosis. Implications for clinical trials. *Archives of Neurology* 1989;46:1112.
49. Patzold U, Pocklington PR. Course of multiple sclerosis: first results of a prospective study carried out on 102 MS patients from 1976-1980. *Acta Neurologica Scandinavica* 1982;65:248-66.
50. Patwardhan M, Matchar D, Samsa G, McCrory D, Li T, Williams R, *et al.* Cost of multiple sclerosis by level of disability: A review of literature. Abstract presented at ISHTAC 2003, Canmore, Alberta.
51. Kobelt G, Lindgren P, Parkin D, Francis DA, Johnson M, Bates D, and *et al.* Costs and quality of life in multiple sclerosis. A cross-sectional observational study in the UK. *Scandinavian Working Papers in Economics*, 2000. Available from: <http://www.swopec.hhs.se/hastef/papers/hastef0398.pdf> [Accessed 12 May 2001]
52. Grima DT, Torrance GW, Francis G. Cost and health related quality of life consequences of multiple sclerosis. *Multiple Sclerosis* 2000;6:91-8.
53. Stolp-Smith KA, Atkinson EJ, Champion ME, O'Brien PC, Rodriguez M. Health care utilization in multiple sclerosis: a population-based study in Olmsted County, MN. *Neurology* 1998;56 :1594-600.
54. Bourdette DN, Prochazka AV, Mitchell W, Licari P, Burks J, and the VA Multiple Sclerosis Rehabilitation Study Group. Health care costs of veterans with multiple sclerosis: implications for the rehabilitation of MS. *Archives of Physical Medicine and Rehabilitation* 1993;74:26-31.
55. ABT Associates, Inc. Cost of resource utilization from the Sonya Slifka Longitudinal Multiple Sclerosis Study for modeling cost effectiveness of disease modifying therapies. 2005.
56. Metlife. The Metlife market survey of nursing home and home care costs. 2005. Available from: <http://www.metlife.com/> [Accessed 05/11/2005]
57. O'Brien, J, Ward AJ, Patrick AR, Caro J. Cost of managing an episode of relapse in multiple sclerosis in the United States. *BMC Health Services Research* 2003;3.

58. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.
59. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technology Assessment* 1999;3:i-134.
60. Claxton K. The irrelevance of inference: A decision-making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* 1999;18:341-64.
61. O'Hagan A, McCabe C, Akehurst R, Brennan A, Briggs A, Claxton K *et al.* Incorporation of uncertainty in health economic modelling studies. *Pharmacoeconomics* 2005;23:529-36.
62. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: The role of cost-effectiveness acceptability curves. *Health Economics* 2001;10:779-87.
63. Kobelt G, Berg J, Atherley D, Hadjimichael O, and Jönsson B. Costs and quality of life in multiple sclerosis A cross-sectional study in the USA. *Scandinavian Working Papers in Health Economics (S-wOPEC)*. No 594. 2005.
64. Rudick RA, Goodkin DE, Jacobs LD, Cookfair DL, Herndon RM, Richert JR *et al.* Impact of interferon beta-1a on neurologic disability in relapsing multiple sclerosis. *Neurology* 1997;49:358-63.
65. Sandberg-Wollheim M, Hommes OR, Hughes RA, Paty DW, Abdul-Ahad AK. Recombinant human interferon beta in the treatment of relapsing-remitting and secondary progressive multiple sclerosis. *Multiple Sclerosis* 1995;1 Suppl 1:S48-S50.
66. Secondary Progressive Efficacy Clinical Trial of Recombinant Interferon-beta-1a in MS (SPECTRIMS). Randomized controlled trial of interferon- beta-1a in secondary progressive MS: Clinical results. *Neurology* 2001;56:1496-504.
67. Cohen JA, Cutter GR, Fischer JS, Goodman AD, Heidenreich FR, Kooijmans MF *et al.* Benefit of interferon beta-1a on MSFC progression in secondary progressive MS. *Neurology* 2002;59:679-87.
68. Fernandez O. A randomized multicentric study on the effects of natural beta interferon treatment in multiple sclerosis during 'relapsing- remitting' with nuclear magnetic resonance. *Nuova Rivista Di Neurologia* 1995;5:3-4.
69. Patti F, L'Episcopo MR, Cataldi ML, Reggio A. Natural interferon-beta treatment of relapsing-remitting and secondary-progressive multiple sclerosis patients. A two-year study. *Acta Neurologica Scandinavica* 1999;100:283-9.

70. Polman C, Barkhof F, Kappos L, Pozzilli C, Sandbrink R, Dahlke F et al. Oral interferon beta-1a in relapsing-remitting multiple sclerosis: a double-blind randomized study. *Multiple Sclerosis* 2003;9:342-8.
71. Bornstein MB. A placebo-controlled, double-blind, randomized, two-center, pilot trial of Cop 1 in chronic progressive multiple sclerosis. *Neurology* 1991;41:533-9.
72. Pozzilli C. An open randomised trial with two different doses of recombinant interferon beta 1a in relapsing-remitting multiple sclerosis: Clinical and MRI results. *Journal of Neurology* 1997;244:S25.
73. European Study Group on Interferon-Beta-1a in MS. Double-blind randomized multicenter dose-comparison study of interferon-beta-1a (AVONEX): rationale, design and baseline data. *Multiple Sclerosis* 2001;7:179-83.
74. Durelli L, Oggero A, Verdun E, Barbero P, Pipieri A, Isoardo G *et al.* Interferon-beta dose and efficacy: the OPTIMS study. *Neurological Sciences* 2001; 22:201-3.
75. Rudge P, Miller D, Crimlisk H, Thorpe J. Does interferon beta cause initial exacerbation of multiple sclerosis? *Lancet* 1995;345:580.
76. Itoyama Y, Saida T, Tashiro K. Japanese multicenter, randomized, double-blind trial of interferon beta-1b in relapsing-remitting multiple sclerosis: two-year results. *Annals of Neurology* 2000;48 Sept:487.
77. Leary SM, Stevens LN. Study design of a double-blind, placebo-controlled trial of two doses of interferon beta-1b in primary progressive multiple sclerosis. *Multiple Sclerosis* 1997;3 Suppl:337.
78. Ozakbas S. Comparison of interferon beta-1a and interferon beta-1b in relapsing remitting multiple sclerosis: MRI activity and clinical results. *Multiple Sclerosis* 1998;4:377.
79. Patti F, Amato MP, Filippi M, Gallo P, Trojano M, Comi GC. A double blind, placebo-controlled, phase II, add-on study of cyclophosphamide (CTX) for 24 months in patients affected by multiple sclerosis on a background therapy with interferon-beta study denomination: CYCLIN. *Journal of the Neurological Sciences* 2004;223:69-71.
80. Bornstein MB. A pilot trial of Cop 1 in exacerbating-remitting multiple sclerosis. *New England Journal of Medicine* 1987;317:408-14.

81. Jacobs L. Intrathecally administered natural human fibroblast interferon reduces exacerbations of multiple sclerosis: Results of a multicenter, double-blinded study. *Archives of Neurology* 1987;44:589-95.
82. Koch-Henriksen N, Sorensen PS. The Danish national project of interferon-beta treatment in relapsing-remitting multiple sclerosis. *Multiple Sclerosis* 2000;6:172-5.
83. Knobler RL, Greenstein JI, Johnson KP, Lublin FD, Panitch HS, Conway. Systemic recombinant human interferon-beta treatment of relapsing-remitting multiple sclerosis: pilot study analysis and six-year follow-up. *Journal of Interferon Research* 1993;13:333-40.
84. Milanese C, Salmaggi A, La Mantia L, Campi A, Eoli M, Savoiaro M et al. Double blind study of intrathecal beta-interferon in multiple sclerosis: clinical and laboratory results. *Journal of Neurology, Neurosurgery and Psychiatry* 1990;53:554-7.
85. Once Weekly Interferon for MS Study Group (OWIMS). Evidence of interferon beta-1a dose response in relapsing-remitting MS: The OWIMS study. *Neurology* 1999;53:679-86.
86. Galetta SL. The controlled high risk Avonex multiple sclerosis trial (CHAMPS Study). *Journal of Neuro-Ophthalmology* 2001;21:292-5.
87. Comi G. Early treatment of multiple sclerosis with Rebif (recombinant human interferon beta): design of the study. *Multiple Sclerosis* 1995;1 Suppl 1:S24-S27.
88. Milanese C, La Mantia L, Palumbo R, Martinelli V, Murialdo A, Zaffaroni M et al. A post-marketing study on interferon beta 1b and 1a treatment in relapsing-remitting multiple sclerosis: different response in drop-outs and treated patients. *Journal of Neurology, Neurosurgery and Psychiatry* 2003;74:1689-92.
89. PRISMS Study Group and the University of British Columbia MS/MRI Analysis Group. PRISMS-4: Long-term efficacy of interferon-beta-1a in relapsing MS. *Neurology* 2001;56:1628-36.
90. Flechter S, Vardi J, Pollak L, Rabey JM. Comparison of glatiramer acetate (Copaxone) and interferon beta-1b (Betaferon) in multiple sclerosis patients: an open-label 2-year follow-up. *Journal of the Neurological Sciences* 2002;197:51-5.
91. Khan OA, Tselis AC, Kamholz JA, Garbern JY, Lewis RA, Lisak RP. A prospective, open-label treatment trial to compare the effect of IFNbeta-1a (Avonex), IFNbeta-1b (Betaseron), and glatiramer acetate (Copaxone) on the relapse rate in relapsing--remitting multiple sclerosis: results after 18 months of therapy. *Multiple Sclerosis* 2001;7:349-53.

92. Wolinsky JS, Comi G, Filippi M, Ladkani D, Kadosh S, Shifroni G *et al.* Copaxone's effect on MRI-monitored disease in relapsing MS is reproducible and sustained. *Neurology* 2002;59:1284-6.
93. Fernandez O, Arbizu T, Izquierdo G, Martinez-Yelamos A, Gata JM, Luque G *et al.* Clinical benefits of interferon beta-1a in relapsing-remitting MS: a phase IV study. *Acta Neurologica Scandinavica* 2003;107:7-11.
94. Murray TJ. Rationale and design of the prospective and retrospective observational study of Avonex and Rebif (PROOF) for the treatment of relapsing-remitting multiple sclerosis. *Current Medical Research & Opinion* 2004;20:25-30.
95. Cook SD. Advancing treatment with interferon beta-1b (Betaferon/Betaseron) in the next decade--thinking beyond the standard dose. *Journal of Neurology* 2003;250 Suppl 4:IV15-IV20.
96. Bencsik K, Rajda C, Fuvesi J, Jardanhazy T, Torok M, Vecsei L. Experiences with interferon-beta-1b treatment in MS after three year follow-up. *Swiss Medical Weekly* 2002;132:237.