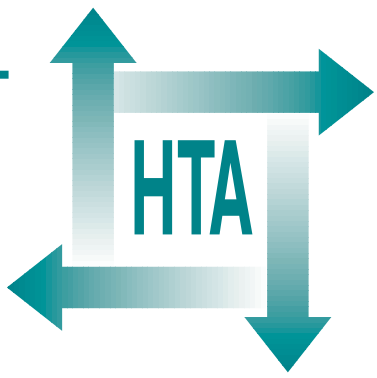
Health Technology Assessment 1998; Vol. 2: No. 4

A cost–utility analysis of interferon beta for multiple sclerosis

D Parkin
P McNamee
A Jacoby
P Miller
S Thomas
D Bates

Health Technology Assessment
NHS R&D HTA Programme



Standing Group on Health Technology

Chair: Professor Sir Miles Irving,
Professor of Surgery, University of Manchester, Hope Hospital, Salford †

Dr Sheila Adam,
Department of Health

Professor Martin Buxton,
Professor of Economics, Brunel University †

Professor Angela Coulter,
Director, King's Fund, London

Professor Anthony Culyer,
Deputy Vice-Chancellor, University of York

Dr Peter Doyle,
Executive Director, Zeneca Ltd,
ACOST Committee on Medical Research & Health

Professor John Farndon,
Professor of Surgery, University of Bristol †

Professor Charles Florey,
Department of Epidemiology & Public Health, Ninewells Hospital & Medical School, University of Dundee †

Professor John Gabbay,
Director, Wessex Institute for Health Research & Development †

Dr Tony Hope,
The Medical School, University of Oxford †

Professor Howard Glennester,
Professor of Social Science & Administration, London School of Economics & Political Science

Professor Sir John Grimley Evans,
Department of Geriatric Medicine, Radcliffe Infirmary, Oxford †

Mr John H James,
Chief Executive, Kensington, Chelsea & Westminster Health Authority

Professor Richard Lilford,
Regional Director, R&D, West Midlands †

Professor Michael Maisey,
Professor of Radiological Sciences, UMDS, London

Dr Jeremy Metters,
Deputy Chief Medical Officer, Department of Health †

Mrs Gloria Oates,
Chief Executive, Oldham NHS Trust

Dr George Poste,
Chief Science & Technology Officer, SmithKline Beecham †

Professor Michael Rawlins,
Wolfson Unit of Clinical Pharmacology, University of Newcastle-upon-Tyne

Professor Martin Roland,
Professor of General Practice, University of Manchester

Mr Hugh Ross,
Chief Executive, The United Bristol Healthcare NHS Trust †

Professor Ian Russell,
Department of Health, Sciences & Clinical Evaluation, University of York

Professor Trevor Sheldon,
Director, NHS Centre for Reviews & Dissemination, University of York †

Professor Mike Smith,
Director, The Research School of Medicine, University of Leeds †

Dr Charles Swan,
Consultant Gastroenterologist, North Staffordshire Royal Infirmary

Dr John Tripp,
Department of Child Health, Royal Devon & Exeter Healthcare NHS Trust †

Professor Tom Walley,
Department of Pharmacological Therapeutics, University of Liverpool †

Dr Julie Woodin,
Chief Executive, Nottingham Health Authority †

† Current members

HTA Commissioning Board

Chair: Professor Charles Florey, Department of Epidemiology & Public Health,
Ninewells Hospital & Medical School, University of Dundee †

Professor Ian Russell,
Department of Health, Sciences & Clinical Evaluation, University of York *

Dr Doug Altman,
Director, Institute of Health Sciences, Oxford †

Mr Peter Bower,
Independent Management Consultant, Newcastle-upon-Tyne †

Ms Christine Clark,
Hon. Research Pharmacist, Hope Hospital, Salford†

Professor David Cohen,
Professor of Health Economics, University of Glamorgan

Mr Barrie Dowdeswell,
Chief Executive, Royal Victoria Infirmary, Newcastle-upon-Tyne

Professor Martin Eccles,
Professor of Clinical Effectiveness, University of Newcastle-upon-Tyne †

Dr Mike Gill,
Brent & Harrow Health Authority †

Dr Jenny Hewison,
Senior Lecturer, Department of Psychology, University of Leeds †

Dr Michael Horlington,
Head of Corporate Licensing, Smith & Nephew Group Research Centre

Professor Sir Miles Irving
(Programme Director), Professor of Surgery, University of Manchester, Hope Hospital, Salford †

Professor Alison Kitson,
Director, Royal College of Nursing Institute †

Professor Martin Knapp,
Director, Personal Social Services Research Unit, London School of Economics & Political Science

Dr Donna Lamping,
London School of Hygiene & Tropical Medicine †

Professor Theresa Marteau,
Director, Psychology & Genetics Research Group, UMDS, London

Professor Alan Maynard,
Professor of Economics, University of York †

Professor Sally McIntyre,
MRC Medical Sociology Unit, Glasgow

Professor Jon Nicholl,
Director, Medical Care Research Unit, University of Sheffield †

Professor Gillian Parker,
Nuffield Professor of Community Care, University of Leicester †

Dr Tim Peters,
Department of Social Medicine, University of Bristol †

Professor David Sackett,
Centre for Evidence Based Medicine, Oxford

Professor Martin Severs,
Professor in Elderly Health Care, Portsmouth University †

Dr David Spiegelhalter,
MRC Biostatistics Unit, Institute of Public Health, Cambridge

Dr Ala Szczepura,
Director, Centre for Health Services Studies, University of Warwick †

Professor Graham Watt,
Department of General Practice, Woodside Health Centre, Glasgow †

Professor David Williams,
Department of Clinical Engineering, University of Liverpool

Dr Mark Williams,
Public Health Physician, Bristol

Dr Jeremy Wyatt,
Institute for Health Sciences, University College London †

* Previous Chair
† Current members

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

A cost–utility analysis of interferon beta for multiple sclerosis

D Parkin[1]
P McNamee[1]
A Jacoby[1]
P Miller[1]
S Thomas[2]
D Bates[3]

[1] School of Health Sciences, University of Newcastle-upon-Tyne, UK
[2] Department of Pharmacological Sciences, University of-Newcastle-upon-Tyne, UK
[3] Department of Clinical Neurology, University of Newcastle-upon-Tyne, UK

Published April 1998
Update (October 1999) see page 57

This report should be referenced as follows:

Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D. A cost–utility analysis of interferon beta for multiple sclerosis. *Health Technol Assessment* 1998; **2**(4).

Health Technology Assessment is indexed in Index Medicus/Medline and Excerpta Medica/Embase. Copies of the Executive Summaries are available from the NCCHTA web site (see overleaf).

NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Pharmaceutical Panel.

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will, in England, be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein
Assistant Editor: Jane Robertson

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 1703 595 639 Email: hta@soton.ac.uk
http://www.soton.ac.uk/~hta

Contents

Glossary and list of abbreviations

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context but a glossary is provided for the non-specialist reader. In some cases usage differs in the literature but the term has a constant meaning throughout this review.

Glossary

Cost-effectiveness ratio The cost of achieving a given level of effect from a healthcare intervention. In this report, cost-effectiveness ratio is expressed as the cost of avoiding one relapse.

Cost–utility ratio The cost of achieving a given level of utility gain from a health-care intervention. In this report, cost–utility ratio is expressed as the cost of gaining one quality-adjusted life year.

EQ-5D The EuroQoL five dimension scale for describing health states.

MSQOL-54 The Multiple Sclerosis Quality of Life 54-item health survey questionnaire.

Quality-adjusted life year (QALY) A measure standardised to one-year periods, which combines levels of health (measured in terms of their impact upon quality of life), with their duration.

SF-36 The Short Form 36-item health survey questionnaire, on which MSQOL-54 is based.

Utility The value given to states or the results of actions. In this report, utility is the value of health states (measured as quality of life levels), or the consequences of healthcare interventions (measured in terms of QALYs).

York MVH tariff A list of values for EQ-5D health states constructed by the Measuring and Valuing Health (MVH) project at York University.

List of abbreviations

CCOHTA	Canadian Coordinating Office for Health Technology Assessment
CEA	cost-effectiveness analysis
CUA	cost–utility analysis
EDSS	expanded disability status scale
IFβ-1a	interferon beta-1a
IFβ-1b	interferon beta-1b
MRI	magnetic resonance imaging
MS	multiple sclerosis
QALY	quality-adjusted life year
QOL	quality of life
RRMS	relapse-remitting multiple sclerosis
SD	standard deviation*
TTO	time trade-off

* Used only in tables

Executive summary

Background

The theory that multiple sclerosis (MS) may be caused by an autoimmune response led to experiments with the use of interferons, which are naturally occurring proteins that have immune-modifying properties. Clinical trials have now established that interferon beta preparations do have some effect in reducing MS disease activity. This report details a cost–utility analysis of interferon beta-1b (IFβ-1b) which, at the time this study was commissioned, was the only interferon preparation licensed for use in MS in the UK.

Objectives

- To identify to what extent IFβ-1b generates quality of life (QOL) gains.
- To measure and value QOL gains.
- To assess the net costs to the health service and society associated with IFβ-1b.
- To compare net costs and QOL gains in a cost–utility model.

Methods

Data collection

Data were collected from existing trials of IFβ-1b and from information on the natural history of MS. New data were collected on costs and QOL from a sample of people in England with relapse-remitting multiple sclerosis (RRMS), divided into two groups: those who had had a relapse in the last 6 months (n = 40) and those who had not (n = 62). Half of each group also took part in a utility measurement exercise.

Data analysis using models

The project involved the construction of a cost-effectiveness model for IFβ-1b for RRMS, which was tested using sensitivity analysis.

Main outcome measures

The following outcomes were used in the data analysis:

- differences in QOL between groups of patients with MS and compared with the general population, using the Multiple Sclerosis Quality of Life (MSQOL-54) and EQ-5D measures
- differences in costs between groups of patients with MS
- estimated changes in QOL and costs arising from IFβ-1b therapy
- cost per relapse avoided
- cost per quality-adjusted life year (QALY) gained.

Results

Quality of life

The impact of MS on QOL is measurable.

- Relapse and remission groups both had poorer QOL than the general population either with or without long-standing illness. However, their valuations of health states were higher than those of the general population.
- The relapse group had poorer QOL than the remission group. In addition, the effects of a relapse may continue over several months.
- Worse health states, as identified by the usual MS clinical measure (EDSS), were associated with poorer QOL.
- Few patients experienced symptom-free days over 6 weeks.
- Patients with worse EDSS status reported more symptoms.

Costs

NHS costs were higher in the relapse group than in the remission group, and the higher the EDSS score, the greater the costs in remission. Cost savings due to relapse rate reduction and slower progression associated with the use of IFβ-1b are small compared with its costs.

Cost-effectiveness and cost–utility

Using current information, the best estimate of cost-effectiveness over 5 years was £28,700 per relapse avoided, giving a cost–utility ratio of £809,900 per QALY gained. Allowing for possible, though unconfirmed effects on progression over 5 and 10 years produced cost–utility ratios of £328,300 and £228,300, respectively, per QALY gained. The estimates are robust to changes in assumptions; the most optimistic estimate was £74,500 per QALY gained. Other drug therapies

currently being tested are likely to have levels of cost-effectiveness similar to IFβ-1b.

Conclusions

IFβ-1b produces important occasional short-term gains in QOL to people with RRMS, but these translate into only small gains in QALYs overall. Even with optimistic estimates of longer-term gains the aggregate QALY gains are small. These benefits are achieved only with a large additional cost.

Implications for policy

Prescribing and policy decisions need to be taken on the basis of judgement at a number of different levels, and the results of the study may be helpful to those making these decisions. Patients and clinicians might consider the true extent of the gains in QOL produced by IFβ-1b in the light of the alternatives for improving QOL. Health Authorities have also to consider whether the extra investment required is worthwhile compared with the gains that health care produces for people with other conditions, bearing in mind both efficiency, as indicated by the cost–utility figures, and also equity.

Implications for research

The impact of MS on QOL is substantial and measurable and may not have been measured well by conventional outcome measures. Future studies of MS and of the impact of MS therapies should base outcomes measurement on QOL. Trial data also need to link closely with natural history and cost data. Valuation of problem-specific health utilities in MS is possible and helpful, but raises the issue of which values should be used in an economic evaluation.

Chapter 1

Background

Multiple sclerosis (MS) is a demyelinating disease of the central nervous system. It has debilitating effects, accompanied by neurological symptoms of differing severity, which, over many years, can lead to chronic disability. Four categories of MS have been identified.

- **Benign, or stable.**
- **Relapsing-remitting** (RRMS). This is the most common form of MS, in which the course of the disease is generally stable but there are occasional relapses (or exacerbation of symptoms).
- **Relapsing-progressive.** Each relapse results in progressively greater disability for the MS sufferer.
- **Chronic progressive.**

There is some overlap between these categories, and many people will progress through more than one of them.

Until recently, no specific therapy was available for MS, and patient management consisted of symptom control, provision of physiotherapy and disability aids, and psychiatric and social support. The theory that MS may be caused by an autoimmune response led to experiments with interferons, which are naturally occurring proteins that have immune-modifying properties. Clinical trials have now established that interferon beta preparations do have some effect in reducing MS disease activity.[1] This has led to the licensing of two products in the UK, interferon beta-1b (IFβ-1b) and interferon beta-1a (IFβ-1a).

The role of interferons in clinical practice remains uncertain, however. In addition to the usual problems of using data from trials on selected patients to extrapolate to practice on a general patient population, MS is characterised by a complex relationship between the disease and its impact on people who have it, and the published studies provide no information on the impact of treatment on overall quality of life (QOL). Moreover, in the UK, the cost-effectiveness of interferon therapy has only been explored by illustrative calculations. However, these issues have been investigated in many other countries, and the findings of a similar study commissioned by the Canadian Coordinating Office for Health Technology Assessment (CCOHTA)[2] will be compared with those presented here.

Cost–utility analysis of IFβ-1b in RRMS

The project reported here was commissioned by the NHS Executive's Health Technology Assessment programme as a cost–utility analysis (CUA) of IFβ-1b in RRMS. (When the study was commissioned, only IFβ-1b was licensed for this indication.)

The overall aims of the project were to provide information on QOL and to evaluate the cost-effectiveness of IFβ-1b for RRMS. In order to do this the following objectives were specified:

- to identify whether, and to what extent, IFβ-1b generates additional QOL in comparison with current practice
- to measure and value additional QOL in terms of utility-based preferences for changes in health states
- to assess the net costs to the health service associated with IFβ-1b treatment
- to compare net costs and QOL gains in a cost–utility model that includes sensitivity analysis.

It should be emphasised that the project did not undertake a new trial of IFβ-1b. Instead, it used existing data from trials and information on the natural history of MS, and new data on costs and QOL, which were collected within the project. These data were used to construct a model from which the cost-effectiveness results could be calculated. The aim was to make the model as rigorous and inclusive as possible. However, it was recognised that in constructing a summary cost-effectiveness model, there is a danger that important aspects and details of costs and benefits may be missed. Information on the QOL of people with MS and the likely impact of IFβ-1b on their QOL have not been studied in detail elsewhere. As a result, this study involved a more detailed look at these aspects, exceeding that strictly required for a cost–utility study.

This report discusses the three main elements of the study – measurement of QOL, measurement of costs, and economic evaluation. Three sets of data were collected covering QOL, health state values and costs, and these were brought together by a cost-effectiveness model in order to determine the impact of IFβ-1b therapy on both patients and the NHS.

Chapter 2

Assessment of quality of life

Key issues in the quality of life measurement debate

Measuring the outcomes of medical interventions by means of clinical measures only is now considered of limited value to clinicians, the health service and the wider society. Assessments of effectiveness need to include wider measures of benefits to patients, and particularly those that measure the impact from the patient's point of view. In addition, interventions are evaluated not just by their efficacy but also their cost-effectiveness. As a result, increasing emphasis has been given to the assessment of health status, well-being and QOL.[3–6]

There are a number of important conceptual and methodological issues in assessing QOL in people who are chronically ill, not least of which is the question of how QOL can best be defined. Recent attempts to conceptualise QOL have resulted in the development of a functional definition that is measurable, evaluative over time, subjective, and incorporates five broad domains – physical, occupational, psychological, social and somatic.[7] QOL assessments can thus be seen as providing quantified information about the degree to which a chronic condition and its treatment are perceived by the patient as either enhancing or detracting from their ability to function across these various domains at different stages in their illness. Within clinical settings it may not always be appropriate to measure all domains; deciding precisely which aspects of QOL to measure will depend upon a number of factors including the nature of the population and condition under study, the source of the data, the predicted costs and benefits of treatment, and the length of the observation time.[8] The choice will also depend to some degree on the availability of suitable instruments and the environment in which the measurement will be conducted.[8] Methodological questions in QOL assessment include:

- From whom should the data be collected?
- Which instrument is most appropriate to measure QOL?
- Is the selected instrument psychometrically sound?

There are also a number of practical considerations in making the assessment. The earliest QOL measures were designed to be completed by physicians or other health professionals. However, findings from studies comparing patient ratings of their QOL with those of healthcare providers and significant others have shown there is limited agreement between individual patients and their observers, with least agreement for QOL phenomena which are difficult to observe.[8] As a result, it is now generally agreed that judgements about the way illness and its treatment affects QOL should be made by patients themselves wherever possible.

The question of whether to use QOL instruments that are generic or those which are disease-specific has been considered by a number of authors.[9–13] There are many advantages of generic measures: they are designed to cover the complete spectra of function, disability and distress relevant to QOL; they are applicable across different types and severities of disease and different medical treatments or health interventions and so allow comparisons to be made across different populations and clinical conditions; and they have established psychometric properties. However, disease-specific measures may be more sensitive to particular problems and outcomes associated with specific conditions and treatments, and have been shown to be useful in clinical trials for a range of conditions. There are strengths and weaknesses of both generic and disease-specific measures, and it has been suggested that the most satisfactory approach to QOL assessment is to use a standard core instrument with customised additions, depending on the particular problem and setting.[8] The basic requirements of both generic and disease-specific measures[14,15] are as follows:

- **validity** (i.e. measure what they purport to measure and relate to other variables in previously hypothesised ways)
- **reliability** (i.e. produce the same results on repeated occasions under similar test conditions)
- **responsiveness** (i.e. detect clinically significant within-patient changes over time); this requirement is the most crucial to clinical trials and evaluation research.

Some researchers have taken issue with the standard approach to QOL assessment and have tried to develop a less structured, patient-elicited approach.[16–18] Unfortunately, this approach requires considerable resources, in terms of both time and interviewer training, making it difficult to incorporate into clinical trials. Another recent development has been the use of health diaries to measure QOL.[19–21] The advantages of health diaries[22] are that they minimise recall error and memory lapse, and are as good as interviews for counting chronic illness and health services use. In addition, diaries are better for recording health problems that are transient and of low enough impact not to prompt medical attention. They may therefore be particularly useful for outcomes that vary from day to day, and in situations where charting the patient's fluctuating experience over time is important.[23] Limitations of health diaries are that they may be subject to 'first-day effects' (i.e. better completion on the first day and during the first week), the data will be biased towards competent diary keepers, and methods of analysis for diary data are, as yet, relatively under-developed.[24]

Assessment of outcome in multiple sclerosis

Traditionally, outcome in MS has been assessed using clinician-based measures that focus largely on impairment. Recently, greater recognition of the relevance and validity of patients' own outcome assessments in this and other clinical areas has stimulated interest in both the application of existing generic health status measures and the development of novel condition-specific ones. The latter have been developed with the aid of qualitative and quantitative studies, which have examined MS patients' subjective experiences of the disease.[25–32] These studies have highlighted the problems of coping with uncertainty, managing treatment regimens, dealing with a changing body and changed self-image and self-esteem, the impact on personal relationships and the organisation of everyday life.

Clinician-based measures

The best-known and most used are those developed by Kurtzke for rating neurological impairment in MS. Others include the Scripps Neurological Rating Scale (SNRS), the Cambridge Multiple Sclerosis Basic Score (CAMBS) and the Guy's Neurological Disability Scale (GNDS).

Kurtzke Functional Systems

Kurtzke Functional Systems consist of eight functional categories: pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, cerebral mental, and spasticity.[33] For each category, scores range from 0 to 6, with high scores indicating greater impairment.

The Kurtzke Expanded Disability Status Scale (EDSS)

This is a global rating of neurological impairment consisting of 20 statements that describe incremental reductions in function.[33] The global rating score ranges from 0 to 10, with normal function (score of 0) progressing through signs and symptoms, problems with mobility, upper limb and bulbar functions, and resulting in death due to MS (score of 10). In relation to both its reliability and responsiveness, EDSS has been subject to considerable criticism.[34–36] One other area of dissent concerns the significance of a change in EDSS score; Kurtzke defined patient improvement or deterioration as a change of one point,[37] but Amato *et al.* suggest that so small a change may be clinically insignificant and argue the need for a two-point difference.[38] Recently a simplified version of EDSS has been developed,[39] from which a patient-rated evaluation has been constructed for use in community studies where individual clinical examination is impossible.[40]

The SNRS

This rating scale grades impairment using a standard neurological examination with additional categories for bowel, bladder, and sexual dysfunction.[41] Scores range from 0 (maximum abnormality) to 100 (normal). Though the authors provide evidence of reliability, its validity and responsiveness have not been shown.

The CAMBS

CAMBS assesses clinical status in relation to impairment, disability, relapse, disease progression and handicap, each being rated on a five-point scale where 1 is best and 5 worst.[42] CAMBS has the advantage over EDSS and SNRS that it can be completed by a trained interviewer without the need for formal examination by a clinician, though to date, only limited work has been done to validate this assessment.

The GNDS

This scale assesses disability across 12 different categories – mental, mood, visual, speech, swallowing, arm and leg function, bowel and bladder function, sexual function, fatigue and other disability (such as pain, dizziness). Severity of each type of disability is graded from 0 (absence of disability) to 6 (total loss of function) by a trained assessor, who scores patient or carer responses to

sequential questions.[43] The authors give evidence of the reliability and validity of GNDS and are currently examining its sensitivity to change. A patient-completed version is also being developed.

Patient-based measures

Patient-based measures for assessing outcome of MS are currently limited, only one having been developed entirely from first principles. The others consist of an existing measure as a core, to which additional MS-specific items have been appended.

The Functional Assessment of Multiple Sclerosis Quality of Life Instrument (FAMS)

FAMS is a self-report measure consisting of 44 items within six domains – mobility, symptoms, emotional well-being, general contentment, thinking/fatigue, and family/social well-being. A further 15 domains cover other concerns including treatment side-effects and sexual function.[44] It was developed as an extension to the Functional Assessment of Cancer Therapy Instrument[45] so that many of the items are not MS-specific, though those that are were generated by patients, providers and a literature review. Although scores on the mobility sub-scale were highly correlated with both EDSS and SNRS scores in the validation exercise, scores on the other sub-scales were not, indicating that FAMS measures aspects of life quality not captured by existing clinician-based measures. The authors provide evidence of the reliability and validity of this measure and suggest FAMS is appropriate for use in clinical trials, though no evidence of its responsiveness is given.

The Leeds Multiple Sclerosis Quality of Life Instrument (Leeds MSQoL)

This QOL measure is a 16-item self-report measure developed *de novo* from focus groups with MS patients.[46] The measure uses a four-point Likert-type response where scores on each item are summed to produce an overall score, with higher scores representing better QOL. Analysis of the structure of the instrument confirmed the presence of a single underlying construct, which was surprising given the emphasis in the literature on the multidimensional nature of QOL. Preliminary evidence suggests that the scale is valid and reliable, and the authors are now planning to apply it in a community-based study of QOL in a prevalent population of 900 people with MS. A potential problem with the Leeds scale is that, although its authors claim it is condition-specific, the final 16 items, which were selected after psychometric analysis from a larger pool of 25, are very general in their coverage.

The Multiple Sclerosis Quality of Life Instrument (MSQOL-54)

MSQOL-54 is a 54-item questionnaire comprising a well-validated generic health status measure, the short form 36-item health survey questionnaire (SF-36),[47] and 18 additional items which are condition-specific.[48] The SF-36 addresses eight distinct domains – physical and social function, roles physical and emotional, pain, energy, mental health, and general health. The authors of MSQOL-54 added a further item each to three of these (social function, pain and energy). The remaining 15 novel items cover domains of health distress, sexual function and satisfaction with sexual function, cognitive function and overall QOL. The psychometric properties of MSQOL-54 were examined in a sample of 179 patients. Both internal consistency and test–retest reliability were high and construct validity was supported by significant associations between scale scores and MS severity, level of ambulation, depressive symptoms and hospital admissions in the previous year. Reproducibility of the MSQOL-54 was assessed by product-moment and intraclass correlations, which ranged from 0.66 to 0.96 for 76 subjects who completed a second questionnaire within 30 days of the first. Factor analysis to examine inter-relationships between the 12 scales suggested the presence of two main factors, the first with eight scales loading onto it, the second with five. Based on these findings, the authors also developed two composite scores of physical and mental health. One reservation about the development of MSQOL-54 is that the additional items were constructed by clinicians and a nurse specialist, apparently without any consultation with patients themselves.

Chapter 3

Assessment of costs

Types of costs

The assessment of cost permits explicit consideration of resource consumption in decisions regarding the use of healthcare interventions. It usually falls into two broad categories. First, there are costs directly related to the provision of the intervention and consequent use of scarce healthcare resources. In the case of IFβ-1b, the main cost is drug acquisition. Second, there are resource use consequences of changes in health status and health care. Examples are length of hospital stay, use of competing or complementary treatments or services, and changes in the productive capacity of individuals. Decisions about the measurement and valuation of such costs are aided by the separation of costs into three elements: direct costs, indirect costs and intangible costs.

Direct costs

Direct costs refer to the resources consumed by a healthcare intervention and any associated events. Such costs can fall on the healthcare system and comprise items such as medical time, nursing time, drugs, equipment and supplies. Alternatively, implementation of an intervention may lead to costs incurred by patients and carers, for example, transportation to hospital and time spent caring for patients.

Indirect costs

Indirect costs refer to changes in the productive use of time by patients and others. The most important item in this category is the change in productivity as a result of changes in disability or life expectancy brought about by the healthcare intervention, for example, lost time from work. Other examples include changes in the amount of time available to pursue other activities, such as leisure.

Intangible costs

Intangible costs relate to changes in health status brought about by the healthcare intervention. For example, changes in pain, social functioning, ability to perform activities of daily living and anxiety. Such changes are not usually explicitly valued in monetary terms. Intangible costs can be measured and valued by health state utility or willingness-to-pay methods.

The present study was undertaken from a social perspective, with direct healthcare costs quantified using monetary values and intangible costs valued using health state utilities. Direct non-healthcare costs (e.g. travel time, caregiver time) were not measured, because the patients in this study were mostly ambulatory with EDSS scores of 6 or less, so that these costs were likely to be minimal. Moreover, it is difficult to estimate with accuracy additional time incurred, particularly with respect to caregiver time.[49] Indirect costs were not quantified in monetary values, which assumes that short periods of lost work time do not lead to production losses, while long-term absences from work led to no net production losses due to replacement by other individuals. In addition, it may be argued that such indirect costs occur at the individual level and, as such, they are incorporated into health state utility measurement.

Chapter 4

Methods of economic evaluation

The aim of economic evaluation is to assess the efficiency with which healthcare interventions use limited resources to produce health outputs. All economic evaluations have two common features. First, the costs of healthcare interventions are compared with their consequences. Second, an explicit comparison is made with at least one other alternative. In the context of a new therapy, the alternative is usually current management. There are a number of forms of economic evaluation, but in this study only two were considered: cost-effectiveness analysis (CEA) and CUA. These consider costs in exactly the same way, but differ in the way consequences are measured.

Cost-effectiveness analysis

In CEA, consequences are assessed using observable health indicators, such as relapse rates, disability-free days and symptom-free days. The objective is to determine which alternative produces greater benefit, in terms of reductions in adverse indicators, and which alternative costs least. If no alternative is superior to all others on both cost and effectiveness (benefit) grounds, examination of the incremental cost-effectiveness ratio of the more beneficial alternative is usually undertaken,[50] to give a net effect per pound. In the context of an effective new therapy, this formulation yields information about the additional cost that the therapy produces in order to achieve an additional unit of effect. The ratio is defined as:

$$\text{Cost-effectiveness ratio} = \frac{\text{Net costs}}{\text{Net effectiveness}}$$

where net costs and effectiveness are measured by the difference in costs and effectiveness between the new healthcare intervention *A* and alternative *B*:

$$\text{Net costs} = \text{Total costs A} - \text{Total costs B}$$

$$\text{Net effectiveness} = \text{Total effectiveness A} - \text{Total effectiveness B}$$

Cost–utility analysis

CEA is most useful where there is one dimension, for example, relapse rates against which consequences can be measured. However, it is often the case that healthcare interventions produce changes along several different dimensions. Moreover, interpretation of which intervention is more cost-effective becomes problematic when one intervention is superior to another in some dimensions, but inferior in others. In such cases, a judgement is required about the relative importance of different dimensions.

To account for these varying factors, CUA is often employed, particularly when the primary purpose of a healthcare intervention is improvement in QOL. The objective of CUA is identical to CEA in that it aims to determine which alternative produces greater health benefit per pound spent. However, the main feature of CUA is the measurement of consequences in terms of **utilities**. In this context, utilities (U) refer to individual preferences for particular health states under conditions of uncertainty, and can be summarised by a single score such that $0 < U < 1$. These utilities are combined with information on the duration of health states to calculate quality-adjusted life years (QALYs). The cost–utility ratio is then given by:

$$\text{Cost–utility ratio} = \frac{\text{Net costs}}{\text{Net QALYs}}$$

where net costs and QALYs are defined as:

$$\text{Net costs} = \text{Total costs A} - \text{Total costs B}$$

$$\text{Net QALYs} = \text{Total QALYs A} - \text{Total QALYs B}$$

This report contains estimates of cost-effectiveness and cost–utility ratios, defined in terms of **cost per relapse avoided** and **cost per QALY gained**, respectively.

Modelling

Before economic evaluations are warranted, demonstration of the clinical efficacy of healthcare interventions is required. Controlled clinical trials are recognised as the best source of such evidence. However, they are often not designed with economic factors in mind, and therefore data on key parameters that are likely to affect cost-effectiveness are usually absent. One frequent

omission in randomised controlled trials is information on resource consumption and costs.[51]

An important consideration in estimating cost-effectiveness relates to events occurring outside the timescale of published trials. The claim has been made that interferon beta may slow disease progression, but the published evidence relates only to short-term effects. However, as costs have been shown to be related to EDSS scores,[52] disregard of long-term outcomes may lead to bias in calculating cost-effectiveness. Therefore, a method of assessing the effect of disease progression is needed.

To account for these factors, CEA and CUA can be extended by means of modelling. This means the construction of a mathematical model describing the natural history of the problem, the impact of interventions on the natural history, and the results in terms of costs and outcomes. A number of studies have undertaken modelling to assess the natural history of MS for prognostic purposes,[53,54] but no published economic model has been identified.

One useful technique which is commonly applied in modelling is decision analysis, which involves structuring decisions into several component parts.[55] As decision analytic models make explicit the assumptions upon which estimates are made, the extent to which results depend on particular assumptions can be rigorously assessed using sensitivity analysis. Furthermore, such models can be updated as more data become available.

Chapter 5
Study methods

Overview of study design

The cost-effectiveness model constructed for this study depended on existing clinical trials to provide reliable evidence-based information on clinical outcomes, and therefore was based on the endpoints and outcomes used in those trials.[1,56] However, the trials do not provide information that would be directly useful in a CUA. For example, there are no relevant data on the process of care and use of healthcare resources, QOL or utilities. This study therefore required primary data collection in these three areas, ensuring that these could be linked to the clinical measures used by the trials. An attempt was also made to ensure that the model was as generic as possible, enabling it to consider patients and issues not covered by current trials.

It should again be emphasised that this study was not itself a trial and it was not possible directly to obtain comparative data on patients who were currently receiving IFβ-1b therapy and those who were not. The aim was to provide a baseline for the cost–utility model rather than a true comparison between different groups of patients. The patient population from which the sample was collected included those patients within the catchment area of the Neurology Service at Newcastle-upon-Tyne. New data were collected for two separate groups of people: those who had recently experienced a relapse, referred to as the 'relapse group', and those who had not, referred to as the 'remission group'. Due to the short time scale of the project, patients were chosen as they were identified by a research nurse. There is no reason to believe that this sample of patients is unrepresentative of patients for whom IFβ-1b therapy might be thought appropriate.

The relapse group of 40 patients had all experienced a relapse in the 6 months preceding an agreed date, and provided information about the effects of relapses on use of healthcare resources and on patients' QOL during a relapse. The remission group of 62 patients, who had not had a relapse in the preceding 6 months, provided information about resource use and QOL during remission. For both groups, there was a mixture of prospective and retrospective data collection.

For all patients, resource use during the preceding 6 months was assessed using a patient questionnaire and analysis of patient case notes. Their current QOL was also assessed. The remission group were asked to judge QOL during remission by measuring their current health status over time. However, because it was difficult to identify patients at the start of a relapse and to monitor them during it, the assessment of relapses was retrospective; the relapse group was asked to judge QOL both in remission and during a relapse, by describing their recent relapse in terms of changes in QOL.

A relapse was defined as appearance of a new symptom or worsening of an existing one sufficient to require management in hospital, either as an in-patient or day case. The definition adopted in the published clinical trial of IFβ-1b therapy was "The appearance of a new symptom or worsening of an old symptom, attributable to MS; accompanied by an appropriate new neurological abnormality; lasting at least 24 hours in the absence of fever; and preceded by stability or improvement for at least 30 days."[56] A different definition was adopted in this study because it was not possible to undertake a prospective study in the time available. Only relapses that would be recorded in hospital notes were included, and therefore detailed information on relapses, and in particular magnetic resonance imaging (MRI) data on new neurological abnormalities, could not be obtained. Consequently, the average severity of relapses in this study is likely to be greater than that in the above trial, which would lead the analysis to overestimate the likely benefits of IFβ-1b in terms of costs averted and QOL gained. However, the sensitivity analysis undertaken permits the impact of this on the overall cost-effectiveness figures to be explored.

The resource use and QOL data were translated into costs and utilities by undertaking two further studies. The cost study was straightforward, involving generation of standard unit costs per treatment course, using NHS financial sources. The utility study had several elements. The use of EuroQol[57] as one of the QOL measures for the patients in both the relapse and the remission study (see below) enabled direct calculation of utility scores using the tariffs published by the York Measuring and Valuing Health (MVH) group.[58] This provided

a problem-generic and general population-based measure. To provide problem-specific utilities, utility measurement was undertaken for a sub-sample of 50 patients, drawn from both relapse and remission groups.

The clinical trials of IFβ-1b reported outcomes in terms of changes in EDSS, though this was not the primary outcome. It was, however, the measure most suitable for incorporation into the present QOL study. In order to link the study data to the results of clinical trials, EDSS data for our study population were required. As the EDSS is not routinely collected, estimates were made by the consultant neurologist on the project team.

Assessment of quality of life

Selection of measures

Robinson strongly advocated the use of patient-based measures to enhance clinical information available on patients with MS.[59] Selection of patient-based measures for inclusion in the present cost–utility study was dictated mainly by the short time available, which meant there was no opportunity to develop measures *de novo,* and time to undertake detailed psychometric evaluations of existing instruments was extremely limited.

There are two types of QOL measures. Profiles provide detailed, multi-faceted information and may therefore have considerable potential to detect condition- or treatment-related changes over time. They pose a problem of interpretation, however, if different groups of patients fare better on some sub-scales and worse on others. Index measures, where a single score is derived from aggregation of a number of single items or sub-scale scores, avoid this problem, but may condense information to the point where sensitivity is compromised.

It has been argued that the decision to use index or profile depends partly on the research context:[60] analysis of QOL in order to identify the potential impact of proposed interventions, or to compare the impact of alternative treatments within the framework of a clinical trial is probably best served by use of profiles; whereas indices will be of special value in healthcare resource allocation studies concerned with determining the QOL gains associated with particular treatments.

MSQOL-54

At the start of the study, FAMS had not yet appeared in the academic press and we were unaware of its development. Evidence about the psychometric soundness of the Leeds MSQoL scale was promising, but limited. The MSQOL-54 measure was therefore selected and it had the advantage over the other measures of incorporating SF-36, thus satisfying the general principle,[8] endorsed by Williams[61] in relation to studies of MS patients, to use a generic health status measure alongside MS-specific ones. As MSQOL-54 had been applied only in the USA, the present study provided an opportunity for its further psychometric validation in the UK, where population norms for its stem measure, the SF-36, were already available.

In applying MSQOL-54 to this study, the UK wording for all SF-36 items was used. The wording of the additional items was checked for cross-cultural applicability and considered acceptable in its published form. However, there was some concern over four detailed items of information on sexual function which, it was considered, might be offensive to some patients and so have a negative effect on response rates. Preliminary reports of the performance of MSQOL-54 by its authors also showed that this was the least well completed item and that it was among the least sensitive to measures of MS severity.[48] As lack of time did not permit pilot work to examine formally the acceptability of these items to patients, the team decided, after careful consideration, to omit them from the questionnaire. This meant that sexual function was measured by one item only, which asked patients how satisfied they had been with this over the preceding 4 weeks. The instrument used therefore had fewer than 54 items, but to avoid confusion it is referred to as MSQOL-54 throughout this report.

EuroQol

EuroQol is a generic QOL measure which can be used to generate a single index score. Its *raison d'etre* was "to provide a simple 'abstracting' device for use alongside other more detailed measures of health-related quality of life"[57] and it is recommended that it be used alongside a more detailed condition- or treatment-specific measure and preferably also a comprehensive generic measure that uses a profile approach. It can be used as a descriptive scheme (referred to as EQ-5D), and also in calculating the values of health states. The descriptive classification has five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) in which each dimension has three different levels of severity, producing a total of 243 possible health

states. For valuation purposes, a tariff, representing a single index value for each of the states, can then be applied to the descriptive classification.

In addition, EuroQol contains a visual analogue scale, in the form of a 'thermometer' (where 0 is the worst and 100 is the best imaginable health state), which may be used for direct valuation of health states. As the present study was concerned both with providing descriptive information about QOL of people with MS, and measuring QALY gains associated with treatments for MS, EuroQol was an obvious choice for inclusion.

Health diaries

A potential problem with standard self-report measures administered at specified times is that they may fail to detect and document fluctuations in health states and so present an inaccurate or limited assessment of QOL. Diaries may be a more appropriate method of obtaining such information and have been used successfully to collect information about levels of disability, self-rated health and well-being, and QOL. In previous studies, diary-keeping periods have ranged from 1 week to 1 year but agreement and completion rates appear unrelated to the length of the diary-keeping period.[22] Diaries can be either free-text or structured, with format varying from that of a calendar to that of a questionnaire. Suggested guidelines for structured diaries recommend that they contain a clear set of instructions on how to complete them and a model example of a correctly completed entry, and that each page should cover either a week, a day of the week, or a 24-hour period, or less, depending on the period for completion.[24] Diary completion is known to be better when they are delivered personally, rather than posted.[24]

As evidence suggests that health diaries have the potential for assessing outcome in clinical trials, we decided to develop and test the use of a daily QOL diary alongside a structured questionnaire.

Application of measures

The relapse group completed a structured questionnaire which asked them to describe their recent relapse and how they had been when it was at its worst (EQ-5D) and how they were now (EQ-5D and MSQOL-54 items). The remission group were asked to keep a daily QOL diary and to complete a structured questionnaire that asked how they were now (EQ-5D and MSQOL-54 items). Patients in the remission group completed the questionnaire at the beginning and at the end of the diary-keeping period. The questionnaires were mailed to both relapse and remission patients with a covering letter from the hospital clinician, with up to two 3-weekly reminders to non-responders. In the case of remission patients, Week 1 and Week 6 questionnaires were mailed out some days prior to diary delivery and collection.

A draft version of the diary was piloted in a small number of patients attending the Neurology outpatient clinic. The revised version comprised an A5 spiral-bound booklet covering a 6-week period. Pages were colour-coded so that those relating to different weeks were clearly distinguishable from one another. Although it was not possible to confine the daily entry to a single page, the week and day of the week to which each set of questions referred was clearly indicated throughout. The diaries were hand-delivered at which time the researcher took the patient through a model entry. At the end of the 6-week period they were collected personally. Patients were encouraged to complete the diary at the end of each day, and for each entry were asked to record their activities during that day, what symptoms, if any, they had experienced, how much of a problem the symptoms had been, and the extent to which symptoms interfered with normal activities. In addition, daily health status was assessed by a precoded question and by asking patients to complete a EuroQol-type thermometer. Patients were also asked to indicate whether they had needed help completing the diary entry that day.

Methods for valuing health status

The aim of the utilities study was to produce utility scores specific to RRMS. As actual experience of different health states may alter individuals' valuations of various health states, it is possible that preferences for particular health states may differ between the RRMS population and the general population. Valuations of different health states by patients with RRMS have not previously been measured and new data were therefore collected as part of this study.

A sub-sample of 50 patients (26 from the relapse group and 24 from the remission group) took part in face-to-face interviews that were designed to generate health-state utility values. Patients were invited to take part in the interview by a member of the study team, who telephoned their home to explain the purposes of the interview and to gain verbal consent.

An exercise using time trade-off (TTO) methods was devised. This is a widely used way of eliciting

valuations for a variety of different populations.[62] The general principle is that patients are asked to make a trade-off between a chronic health condition for *t* years and good health for a shorter period of *x* years. The period *x* is varied in length until the patient is indifferent between the two health states. At the point of indifference the valuation *h* of the health condition is given by:

$$h = \frac{x}{t}$$

The exercise consisted of two stages. The first stage required the patient to value, by means of TTO, three different MS-specific health descriptions of a period lasting 3 years, constructed in order to value the main effects of IFβ-1b. They incorporated the effects on QOL, relapses and probability of disease progression, derived from published data on the natural history of MS[63–66] and on the trial effects of IFβ-1b.[1] The scenarios are described in Appendix 1.

The first scenario described moderate QOL effects of MS with no drug therapy. This included a health description in terms of EQ-5D states and involved three relapses requiring treatment in hospital and a certain chance of progression by the end of a 3-year period. It matched the observed outcomes for the placebo group in the published trial, and is referred to as the 'placebo scenario'. The second scenario described the QOL effects of drug therapy designed to alter the natural history of the disease. This included the same EQ-5D-based description. Within the 3-year period this description involved two relapses requiring treatment in hospital, a lower chance of progression, and possible side-effects of therapy. It matched the observed outcomes for the active therapy group in the published trial, and is referred to as the 'IFβ-1b scenario'. (However, the actual name of the drug was not used in the description.) The third scenario described some likely severe QOL effects of MS without drug therapy. In this description more than three relapses would certainly occur and the probability of disease progression was highest. This is referred to as the 'severe scenario'. Respondents were asked to rank the descriptions in order of preference into 'best', 'intermediate' and 'worst' states.

Due to the difficulties of providing an accurate and succinct description of the lifetime course of MS, the scenarios were valued as temporary health states. This involved the respondent being offered choices between a shorter period in a poorer health state (followed by good health) versus a longer period in a better health state (followed by good health). Visual aids were used, which consisted of laminated cards and a board to show the duration of states. Values of *x* for each state were reached by 'ping-ponging' between longer and shorter periods until preferences changed, which is equivalent to indifference between two alternative states.

In order that utilities for the intermediate and best states could be calculated on the conventional scale of 0 = dead, 1 = healthy, the worst state was treated as a chronic state and compared with good health. Specifically, the value for the worst (h_w), intermediate (h_i) and best (h_b) states were found using:

$$h_w = \frac{x}{t}$$

$$h_i = 1 - (1 - h_w)\frac{x}{t}$$

$$h_b = 1 - (1 - h_i)\frac{x}{t}$$

where $t = 3$ and x is the respondent's indifference point.

The second stage of the exercise required patients to value five different EQ-5D health state descriptions. These were selected as the most informative in terms of reflecting different health dimensions and those most commonly encountered by RRMS patients. The values generated from the RRMS population could be directly compared with the published tariffs for the general population provided by the University of York MVH group.[58] (The EQ-5D states chosen had also been measured directly in the MVH study.)

Costing methods

In order to identify differences in healthcare utilisation and costs between the relapse and remission phases of RRMS, the hospital notes of both the remission and the relapse groups were examined. All data on service receipt within the hospital were abstracted for the 6 months prior to the study.

Each patient also completed a postal survey which included questions on the use of a range of health services received within hospital and community settings. This included hospital services, which served simply as a cross-reference for the data extracted from hospital notes. The survey was

the only source of data on the use of community services (general practitioner, home care, district nurse, health visitor, physiotherapy, occupational therapy, speech therapy and social work services). Given the lack of routine community data, the nature of the disease and the large difference in cost between secondary and community services, it was reasonable to rely on patient-supplied data in this area.

Hospital treatment for symptoms of RRMS typically takes place on an inpatient, day case or outpatient basis in medical specialties, where patients are monitored, drugs administered and various tests carried out. Data on three main categories of hospital care were collected: the nature of hospital visits, drugs prescribed, and procedures and tests.

To determine the nature of inpatient visits, data were recorded on specialty, number of admissions and length of hospital stay. For day case and outpatient visits, specialty and number of visits were recorded. The name, dosage and duration of each course of drug therapy were recorded, and all procedures and tests were noted by type and frequency. In addition, the supply of any appliances was recorded. For community services, data were collected by number of visits and the professional group providing the service.

These data built up a profile of 6-months' health service use for each patient. Costing involved the use of several sources, and a list of unit costs for each item of resource use was generated (see Appendix 2). Inpatient costs were based on the Northern & Yorkshire regional average cost per bed day by specialty, from the 1996 Chartered Institute of Public Finance and Accountancy (CIPFA) database.[67] This uses a top down method based on the Trust Financial Returns (TFR1). The bulk of this cost is made up of labour, capital and overhead charges. It does, however, also include an element for drugs, procedures and tests. To calculate costs per patient, the CIPFA cost per bed day by specialty was multiplied by the observed length of stay in each specialty. Day case and outpatient care were costed by multiplying frequency of visits by CIPFA unit cost estimates.

Drug costs were calculated using the British National Formulary.[68] For procedures and tests, costs were provided by the local Trust which provides the regional specialist service for the patients in this study. Costs for appliances and community services were calculated using Personal Social Services Research Unit (PSSRU) unit costs.[69] The cost for IFβ-1b included an amount for administration and monitoring.

Total costs per patient were therefore calculated by summing all inpatient, day case, outpatient, drug, procedure, test, appliance and community costs. This method produces a slight over-estimate because of the inclusion of drugs, procedures and tests in the cost per bed day figures. We therefore calculated an alternative total cost by excluding our data on these items, which produces a slight under-estimate.

This exercise produced a baseline average cost for RRMS patients in remission and for those having experienced a recent relapse. The additional costs associated with relapse could then be estimated by taking the difference between the two estimates.

Decision analytic model

Two models were constructed, both of which aimed to give an indication of the long-term cost-effectiveness of IFβ-1b compared with standard care. The current trial evidence suggests only an effect on the number of relapses. To analyse this, a very simple model was constructed that calculated total costs of treatment and the cost savings and QALY gains from reductions in the number of relapses. However, the IFβ-1b trial also found an effect, which was not statistically significant, on disability progression and it is likely that other trials may demonstrate such a change. (It should be noted that the trial was not constructed with progression as an endpoint.) It was therefore necessary to construct a more complex model to take this into account.

The more complex model was based on a series of health states through which individuals move over time. These states were based on the EDSS because all published randomised controlled trials and epidemiological studies of disease progression use this measure. The model consisted of a hypothetical cohort of patients, within which each patient is at any one time classified into one of the health states. Every patient is initially placed in health state EDSS 3, chosen as the average baseline health state for patients in the published trial of IFβ-1b.[1] Health states in the following years are determined by the probabilities of transition to other EDSS states. The probabilities of avoiding disease progression (EDSS 3→3) were taken from the trial of IFβ-1b[1] and from the natural history literature.[63,64] Probabilities of disease progression

(EDSS 3→4, EDSS 3→5, EDSS 3→6) associated with IFβ-1b therapy were calculated by applying a reduced risk rate to standard care progression probabilities using estimates reported by the trial.

Each transition is associated with particular levels of costs and QOL. These were estimated using the new data collected within this study. For each patient, their self-reported EQ-5D state in remission was converted to a utility score using the 'tariff' values produced by the University of York MVH Group.[58] The patients were then grouped by EDSS state, and an average utility score was calculated for each state. To estimate average EDSS transition costs per year, costs per EDSS state measured over 6 months' costs within the remission group were used. To calculate the average cost of a relapse, the difference in average cost between the remission and relapse group was taken. Transition utilities and costs were then derived by taking a weighted average, dependent on the length of time in each state. All future costs were converted into present values using the Treasury recommended rate of 6%.

The models were estimated over 5 and 10 years. The 5-year model covers the period for which the trial provided evidence. The 10-year model is more speculative, but is relevant to a longer term assessment of the impact on RRMS. The assumptions for the 5- and 10-year models with regard to estimates of probabilities of progression, utilities and costs are summarised in *Table 1*.

As an alternative to the estimation of utilities via the model, the utility scores derived from the TTO exercise were also used. This provided a comparison of two very different sources of preference information. The full model uses population-based values from a generic index applied to events predicted by the model to be certain. The direct method produces patient-based values from a condition-specific measure applied to uncertain events.

TABLE 1 *Assumptions used in base case analysis*

Parameter	Five-year model*		Ten-year model†		Source [Reference number]
	Standard care	IFβ-1b	Standard care	IFβ-1b	
EDSS transition probabilities					
3→3	0.45	0.60	0.18	0.30	[1, 66]
3→4	0.17	0.12	0.21	0.18	[1, 63, 64]
3→5	0.08	0.06	0.11	0.09	[1, 63, 64]
3→6+	0.30	0.22	0.50	0.43	[1, 63, 64]
Years within EDSS states					
3→4					
3	3	3	3	3	[66]
4	2	2	7	7	[66]
3→5					
3	1	1	1	1	[66]
4	2	2	2	2	[66]
5	2	2	7	7	[66]
3→6+					
3	1	1	1	1	[66]
4	1	1	1	1	[66]
5	1	1	1	1	[66]
6	2	2	3	3	[66]
7	0	0	4	4	[66]

continued

TABLE 1 contd *Assumptions used in base case analysis*

Parameter	Five-year model*		Ten-year model†		Source [Reference number]
	Standard care	IFβ-1b	Standard care	IFβ-1b	
EDSS utilities					
3	0.71	0.71	0.71	0.71	Study estimates, [56, 66]
4	0.66	0.66	0.66	0.66	Study estimates, [56, 66]
5	0.52	0.52	0.52	0.52	Study estimates, [56, 66]
6	0.49	0.49	0.49	0.49	Study estimates, [56, 66]
7	–	–	0.35	0.35	Study estimates, [56, 66]
EDSS costs per year (£)					
3	740	740	740	740	Service receipt – own study estimates
4	850	850	850	850	Unit costs – CIPFA, British National
5	1570	1570	1570	1570	Formulary (BNF), Newcastle & North
6	1590	1590	1590	1590	Tyne Health Authority, Royal Victoria
7	–	–	3080	3080	Infirmary, Newcastle-upon-Tyne
Number of relapses	5.23	3.71	9	6	[1]
Relapse length (months)	1	1	1	1	Study estimates
Utility loss per relapse	0.5	0.5	0.5	0.5	Study estimates
Relapse cost (£)	2115	2115	2115	2115	Study estimates
IFβ-1b costs per year (£)	0	10,500	0	10,500	BNF, study estimates

* ***Assumptions of 5-year model***

1. *Probability of EDSS 3→3 in both groups is the Kaplan-Meier estimate.*[1]
2. *Probability of EDSS 3→6 for standard care is estimated from reference 64.*
3. *Probability of EDSS 3→4, EDSS 3→5 for standard care is estimated from reference 63, which shows that twice as many patients were assigned to EDSS 4 compared with EDSS 5.*
4. *Probability of transitions EDSS 3→4, EDSS 3→5, EDSS 3→6 for IFβ-1b is estimated with the equation* pb = pp – [pp × (wp – wb / wp)], *where* pb *is the probability of transition associated with IFβ-1b therapy,* pp *is the probability of transition associated with standard care,* wp *and* wb *are the proportion of standard care and IFβ-1b patients worsening by at least one point on the EDSS instrument, respectively.* wp = 0.55 *and* wb = 0.40 *using estimates published in reference 1.*
5. *QALY gains are calculated using* p[(u × t) – (u^/12 × n)], *where* p = *probability of transition,* u = *remission utility,* t = 5, u^ = *relapse utility, and* n = *number of relapses.*
6. *All future costs are discounted to present values using the Treasury rate of 6%.*

† ***Assumptions of 10-year model (base case)***

1. *Probability of EDSS 3→3 for standard care is estimated from reference 66.*
2. *Probability of EDSS 3→6 for standard care is estimated from reference 64.*
3. *Probability of EDSS 3→4, EDSS 3→5 for standard care is estimated from reference 63, which shows that twice as many patients were assigned to EDSS 4 compared with EDSS 5.*
4. *Probability of transitions EDSS 3→4, EDSS 3→5, EDSS 3→6 for IFβ-1b is estimated with the equation* pb = pp – [pp × 0.5(wp – wb / wp)], *where* pb *is the probability of transition associated with IFβ-1b therapy,* pp *is the probability of transition associated with standard care,* wp *and* wb *are the proportion of standard care and IFβ-1b patients worsening by at least one point on the EDSS instrument, respectively.* wp = 0.55 *and* wb = 0.40 *using estimates published in reference 1, with* pb *(EDSS 3→3) = 1 – pb.*
5. *QALY gains are calculated using* p[(u × t) – (u^/12 × n)], *where* p = *probability of transition,* u = *remission utility,* t = 10, u^ = *relapse utility, and* n = *number of relapses.*
6. *All future costs are discounted to present values using the Treasury rate of 6%.*

Chapter 6

Results – quality of life

Demographic and clinical characteristics

A total of 102 MS patients were included in the study. In keeping with the known epidemiological features of MS,[70] a high proportion of patients in the study were women (72%). The mean age was 42 years (range: 25–65 years); 79% were currently married; only 23% were currently in paid employment, 50% describing themselves as unable to work due to long-term illness or disability. This figure was slightly higher among patients who had recently had a relapse (55%) than among those in remission (47%). In both the relapse and the remission groups, about 10% had been diagnosed in the last 2 years, and about 40% in the last 5 years. A higher percentage of patients in the relapse group had been diagnosed for 10 years or more (38% compared with 28%). Forty per cent of the remission group reported no relapses in the previous year and a further 14% had experienced only one; 19% reported at least three relapses. Among relapse patients 23% had experienced only one and 33% had suffered at least three. EDSS scores were available from the hospital notes for 89 patients of whom 37% had a score of 0–3, 36% of 3.5–5, and 27% of 6 or more.

MSQOL-54 profile in relapse and remission

A detailed analysis of the psychometric properties of MSQOL-54 was provided by application of the MAP-R analysis package developed by Ware and co-workers.[71] The results of this analysis are reproduced in Appendix 3. Scores on the various domains of MSQOL-54 for those in remission and those having had a recent relapse are given in *Table 2*. Mean scale scores ranged from 13.2 on role physical to 63.7 on mental health for those having recently had a relapse, and from 39.9 on energy/vitality to 73.1 on bodily pain for those in remission. There were highly statistically significant differences ($p < 0.001$) between the two groups in scores on the physical function, role physical, and social function scales, in the change in health item and in the physical health composite score. Role limitations due to emotional problems, mental health, cognitive function and general

TABLE 2 *MSQOL-54 scores for patients in remission and relapse groups*

Scale*	Remission group Mean (SD)	Relapse group Mean (SD)	2-tailed *p* value†
Physical function	42.9 (28.5)	20.0 (18.6)	0.0001
Role physical	39.9 (40.3)	13.2 (28.9)	0.0004
Role emotional	64.3 (44.0)	53.5 (47.5)	0.29
Bodily pain	73.1 (23.3)	58.6 (27.0)	0.01
Mental health	69.4 (19.5)	63.7 (22.9)	0.25
Energy/vitality	39.9 (18.9)	32.0 (17.7)	0.05
Health distress	59.2 (26.4)	48.8 (30.5)	0.09
Social function	66.2 (22.7)	47.5 (25.9)	0.0006
Cognitive function	70.1 (25.3)	60.8 (29.7)	0.14
Sexual function	59.6 (32.3)	46.1 (33.7)	0.05
General health	43.2 (21.9)	39.0 (22.3)	0.31
Overall QOL	69.5 (61.6)	49.9 (20.6)	0.008
Change in health	47.9 (22.2)	31.9 (22.6)	0.0007
Physical health composite score	43.9 (16.9)	32.8 (15.5)	0.0009
Mental health composite score	63.5 (23.7)	55.6 (24.6)	0.11

* *For each scale, a small number of patients did not complete sufficient items to permit computation. Bases on which scale scores were calculated range from 53 to 60 for remission patients, and from 38 to 40 for relapse patients.*
† *Mann-Whitney Test.*
SD = standard deviation.

health perceptions were least sensitive to group differences in remission status.

Mean MSQOL-54 scores for subgroups classified by EDSS score (score of 0–3, 3.5–5.5, 6 and over) are shown in *Table 3*. Mean scores across the scales ranged from 39.3 for energy/vitality to 79.3 for overall QOL in those with low EDSS scores, and 12.6 for physical function to 65.2 for bodily pain in those with high EDSS scores. There was a highly significant trend ($p < 0.0001$) in mean scores for physical function, from 54.2 for those with an EDSS score of 3 or under, to 12.6 for those with a score of 6 or more. Mean score differences were also significant for social and sexual function ($p < 0.01$) and for role physical and health distress ($p = 0.01$). For both physical and mental health composite scores, there was a small but significant difference by EDSS score.

SF-36 scores for the study respondents were compared with the normative scores for a UK general population, calculated from a large-scale community study.[72] As measured by the SF-36, the health status of people with MS was markedly lower than that of people without MS *(Table 4)*. Study patients scored 45 points lower on the physical function scale than a general population sample with long-standing illness and 59 points lower than one without. On the role physical scale, they scored 43 points lower than the general population with, and 62 points lower than the one without long-standing illness. The differences were smaller but nonetheless considerable for social function, energy, and general health perceptions. Interestingly, there was little difference between the three groups in scores on the mental health scale.

EQ-5D profile in relapse and remission

Used descriptively, EQ-5D provides simple information about functioning across five dimensions – mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. For each of these, respondents were asked whether they were having no, some or extreme problems. EQ-5D can thus be seen as providing parallel information, though in much reduced form, to the SF-36. In the present study, those who had recently experienced a relapse were asked to complete EQ-5D in relation to how things had been when the relapse was at its worst and how things were for them now. Patients in remission answered questions on their current health at Week 1 and Week 6 *(Table 5)*.

TABLE 3 *MSQOL-54 scores by EDSS score*

Scale*	EDSS score			*p* value†
	0–3 Mean (SD)	3.5–5.5 Mean (SD)	6 and over Mean (SD)	
Physical function	54.2 (29.8)	25.9 (20.9)	12.6 (12.0)	< 0.0001
Role physical	44.8 (43.0)	19.2 (29.1)	15.8 (37.5)	0.01
Role emotional	67.8 (44.0)	55.9 (46.7)	45.6 (48.7)	0.30
Bodily pain	67.7 (23.8)	65.2 (26.1)	65.2 (28.3)	0.92
Mental health	71.6 (17.5)	67.1 (19.5)	59.8 (26.8)	0.27
Energy/vitality	39.3 (17.5)	36.6 (19.7)	35.3 (21.1)	0.84
Health distress	64.3 (24.7)	56.3 (29.9)	39.3 (28.6)	0.01
Social function	65.7 (26.2)	56.9 (23.1)	41.7 (23.3)	0.005
Cognitive function	71.3 (25.9)	67.9 (23.7)	54.3 (32.5)	0.14
Sexual function	69.8 (31.6)	48.3 (27.5)	42.9 (37.2)	0.007
General health	42.1 (21.8)	37.9 (21.5)	38.6 (25.3)	0.75
Overall QOL	79.3 (85.1)	53.6 (19.1)	49.9 (22.9)	0.04
Change in health	48.3 (27.0)	39.8 (20.9)	34.5 (24.3)	0.25
Physical health composite score	45.4 (18.2)	36.1 (15.7)	31.2 (15.9)	0.03
Mental health composite score	67.5 (24.2)	59.3 (22.9)	49.9 (24.5)	0.05

* *For each scale, a small number of patients did not complete sufficient items to permit computation. Bases on which scale scores were calculated range from 26 to 30 for patients with EDSS of 0–3; from 29 to 32 for patients with EDSS of 3.5–5.5; and from 19 to 21 for patients with EDSS of 6 and over.*

† *Kruskal-Wallis one-way analysis of variance (ANOVA).*

TABLE 4 *SF-36 scores for MS patients and the UK general population*

SF-36 domain	MS sample* Mean (SD)	UK general population with long-standing illness Mean (SD)	UK general population with no long-standing illness Mean (SD)
Physical function	33.2 (27.2)	78.3 (23.2)	92.5 (13.4)
Social function	58.3 (25.7)	80.2 (24.8)	91.3 (15.8)
Role physical	29.2 (38.4)	71.9 (38.9)	91.4 (23.2)
Role emotional	59.9 (45.5)	76.3 (36.4)	85.6 (29.3)
Mental health	67.0 (21.0)	69.9 (18.7)	75.4 (16.3)
Energy/vitality	36.7 (18.7)	54.0 (21.1)	64.0 (18.2)
Pain	67.1 (25.7)	69.8 (25.4)	86.3 (17.9)
General health	41.5 (22.0)	60.8 (23.0)	78.8 (15.7)

* *For each scale, a small number of patients did not complete sufficient items to permit computation. Bases on which scale scores were calculated range from 92 to 100.*

TABLE 5 *EQ-5D responses for patients who had recently relapsed and those in remission (percentages)*

	Relapse (n = 57)*		Remission (n = 32)*	
	At worst	Now	Week 1	Week 6
Mobility				
No problems	3	2	26	27
Some problems	67	98	72	71
Confined to bed	30	–	2	2
Self-care				
No problems	16	42	74	83
Some problems	69	58	24	15
Unable to wash/dress	15	–	2	2
Usual activities				
No problems	3	10	25	33
Some problems	47	73	70	62
Unable to perform	50	17	5	5
Pain/discomfort				
No problems	16	15	44	40
Some problems	50	80	56	60
Extreme problems	34	5	–	–
Anxiety/depression				
No problems	19	48	53	62
Some problems	47	47	45	38
Extreme problems	34	5	2	–

* *Numbers quoted are lowest bases on which percentages were calculated.*

For each domain, the pattern of responses was as predicted. Among the relapse group, the QOL profile was significantly poorer during relapse than 'now': for example, 30% of patients were confined to bed when the relapse was at its worst, whereas none were currently confined to bed; and 50% had been unable to perform their usual activities when the relapse was at its worst compared with only 17% currently. Comparing the current profile of relapse and remission patients, 98% of the former group reported some problems with mobility compared with 72% of the latter (Week 1 data); 58% reported some problems with self-care compared with only 24%; and 80% reported some problems with pain/discomfort compared with only 56%. Interestingly, there was little difference in the current profile for anxiety and depression between the two groups. When Week 1 and Week 6 responses were compared for the remission group, the differences, as expected, were small and not statistically significant.

In addition to examining responses for each EQ-5D dimension separately, the distribution of health states represented in the sample was examined. *Table 6* shows the distribution with health states ordered from highest to lowest according to the value attached to them in the York MVH tariff. Twenty-seven out of the possible 243 states were defined by the 96 respondents completing all five items. Forty-two per cent were classified in terms of three states (21221, 21222, 22222). Only 5% were in state 11111, compared with 57% of respondents in a general population survey.[73] A total of 31 of the reported health states included a problem with mobility and 32 included a problem with performing usual activities. The relapse patients, in response to how things were for them 'today' reported 15 health states, in all but one of which mobility was a problem. When the relapse had been at its worst, they reported 21 states in all but one of which mobility was a problem, and in seven of which the mobility problem was severe. The remission patients reported 21 health states, in 13 of which mobility was a problem.

TABLE 6 *Summary EQ-5D health states*

EQ-5D state*	All – today (n = 96)	Relapse – today (n = 40)	Relapse – at worst (n = 30)	Remission (n = 56)
11111	5	1		4
11211	3		1	3
21111	4	1		3
11112	2			2
12111	1			1
21211	6			6
11221	1			1
21212	4	1		3
21121	4	2		2
11122	1			1
22211	4	2		2
21221	12	7		5
11222	1			1
21122	1		1	1
22212			1	
21222	13	5	1	8
22221	7	5	3	2
22222	15	8	2	7
21311	1			1
22312	1	1		
21322	1			1
22321	1	1		
21313			1	
22322	3	3	1	
21231			1	
12223	1			1
22313			1	
22223			1	
22323	1	1	2	
32321			1	
32222			1	
22232	1	1	3	
32322			1	
33322	1			1
32323			1	
22333	1	1	2	
32332			1	
33332			1	
33333			3	

* *EQ-5D notation represents a composite health state denoting the level of severity in each of the five health dimensions. For example, 11112 means no walking problems, no self-care problems, no usual activities problems, no pain/discomfort, moderate anxiety/depression, while 22222 means moderate problems across all dimensions.*

Diary-based quality of life information

The 62 patients in remission were asked to keep the QOL diary daily for a 6-week period. For each daily entry they were asked to indicate whether they had experienced any of a list of six symptoms that day. If they had, they were asked how much of a problem each had been and to what extent it had interfered with what they wanted to do. They were also asked what activities they had done during the day, how good or bad their health had been (using the EuroQoL thermometer), and whether it had been better, worse or about the same as usual.

Quality of diary completion

Complete information was available for 39 patients (63%): one respondent missed 14 out of 42 entries because they went on holiday and forgot to take the diary; three patients missed seven entries for the same reason; 19 missed between one and five entries because of short trips away from home, or they were too unwell, or because illness in another family member for whom they were caring meant they did not have the time. Out of a maximum possible number of entries from all subjects of 2604, 2526 were completed producing a day entry completion rate of 97%. However, in addition to missing whole days, individual items within each daily entry were also missed. Across all entries for which at least some information was recorded, the rate of failure to complete individual items ranged from 2.7 to 5.1%. The highest failure rate was for the section that contained the EuroQol thermometer.

QOL as indicated by diary responses

Overall, the percentage reporting symptom-free days was low: 45 patients (73%) reported at least one symptom every day. The percentage of symptom-free days out of all days for which the question was completed was similarly very low, only 11%. The percentage of reported symptom-free days fell from 19% for those with low EDSS scores to 8% for those with medium scores and to 0%, meaning no symptom-free days, for those with high scores. Patients were also asked whether they had been able to carry on with usual activities, or whether they had been unable to do as much as usual, or stayed at home though not in bed, or stayed at home in bed, or stayed in hospital. On 71% of all possible occasions when they could have done so, they recorded that they carried on with their usual activities. However, the percentage fell from 73% for those with a low EDSS score to only

TABLE 7 *Percentage diary responses by EDSS scores*[*]

	All scores (n = 62)	EDSS 0–3 (n = 27)	EDSS 3.5–5.5 (n = 17)	EDSS 6+ (n = 11)
Problem-free days	11%	19%	8%	0%
Daily activities				
Carried on as usual	71%	73%	73%	54%
Not able to do as much	19%	24%	12%	21%
Stayed at home, not in bed	10%	3%	14%	25%
Stayed in bed	0%	0%	1%	0%
Health				
About same as usual	67%	70%	70%	71%
Better than usual	15%	15%	13%	8%
Worse than usual	18%	15%	17%	1%

[*] *Presents percentages of occasions on which particular responses were given out of all possible occasions for which responses were obtained.*

54% for those with a high score *(Table 7);* conversely, the percentages of occasions patients had to stay at home, including in bed, rose from 3% in those with low EDSS scores to 25% in those with high scores. Although the percentage of occasions on which patients described their health as 'about the same as usual' did not differ by EDSS score, the percentages of occasions on which they described it as 'better' or 'worse than usual' did *(Table 7).*

Finally, responses to the EuroQol thermometer were examined by calculating an individual mean score across all diary days, and hence a mean score for all patients across all diary days. The mean for individual respondents ranged from 28.0 to 100.0 (one person marked the maximum score throughout the entire diary-keeping period). The mean (and standard deviation) for all subjects across all entries was 62.2 (17.2). Mean scores ranged from 66.2 (17.1) in those with an EDSS score of up to 3, to 46.5 (18.1) in those with an EDSS of 6 or more. Minimum scores across all entries ranged from 10.0 to 100.0, and maximum scores ranged from 40.0 to 100.0.

Health state utilities and preferences

The utilities used in the cost-effectiveness models were derived from patients' EDSS scores, their self-reported EQ-5D health states and the York MVH tariff. *Table 8* shows the distribution of EQ-5D scores within different EDSS states and the mean utility values for the EDSS state, which were derived from the tariff applied to the EQ-5D states. *Table 9* shows the EQ-5D states in remission and relapse, and the tariff values applied to them. From this, the mean value of the utility loss from a relapse was calculated as 0.468 (= 0.604 – 0.136).

TABLE 8 *EQ-5D states and mean tariff (for each EDSS state) from York MVH used to construct values for EDSS utilities*

EDSS	EQ-5D	Mean tariff value
3	12111 11211 21111 11122 21121 21221 21212 21222 22222 (× 2)	0.71
4	21111 21121 21221 (× 3) 21212 21222 22222 (× 3)	0.66
5	11211 21111 22221 (× 2) 21222 22222 (× 3) 21322 33322	0.52
6	22211 (× 2) 21212 21221 21222 21311 22221 22312 22222 (× 3) 22322 22333	0.49
7	22221 22222 22321 22323	0.35

TABLE 9 *EQ-5D states and the mean tariff values from York MVH used to construct value for relapse utility*

EQ-5D state – relapse	Tariff value	EQ-5D state – remission	Tariff value
33333	–0.594	22211	0.710
22232	–0.016	21221	0.691
22232	–0.016	21222	0.620
22221	0.587	22221	0.587
32322	–0.056	22222	0.516
22222	0.516	21221	0.691
32323	–0.221	22221	0.587
22323	0.024	22222	0.516
33333	–0.594	22322	0.189
32332	–0.319	22321	0.260
22322	0.189	22222	0.516
22313	0.147	22221	0.587
21222	0.620	21111	0.850
21231	0.159	21121	0.727
22222	0.516	21221	0.691
32321	0.015	21221	0.691
32222	0.002	21221	0.691
22222	0.516	21221	0.691
22221	0.587	22221	0.691
11211	0.883	11111	1.000
33333	–0.594	22322	0.189
22212	0.639	22211	0.710
Mean value	**0.136**		**0.604**

Four respondents did not complete all items on EQ-5D classification

TABLE 10 *Background data of respondents who participated in utility interviews*

	Number of patients (%)
Type of respondent	
Remission	24 (48)
Relapse	26 (52)
Demography	
Median age (quartiles)	42 (35–49)
Male	17 (34)
Female	33 (66)
Employment status	
Employed	11 (22)
Unable to work	27 (54)
Other	12 (24)
Disease duration and status	
Less than or equal to 5 years	17 (33)
Greater than 5 years but less than or equal to 10 years	17 (33)
Greater than 10 years	16 (33)
Median EDSS (quartiles)	4.0 (2.5–5.5)

For the direct measurement of utilities, everyone who was invited to take part agreed to be interviewed. Three interviews were incomplete or invalid, due to the respondent being unable to complete all valuation tasks. *Table 10* presents background information on those patients who completed the interview. There were almost equal proportions of people in remission and having had a recent relapse. The average age was 42 years and 66% were women. Almost one-quarter were employed, but the majority were unable to work due to functional problems caused by MS. There was an even spread of disease longevity, and the majority of patients were ambulatory with an average EDSS score of 4.0.

Table 11 provides information on the ranking properties of the health state scenarios and the EQ-5D states. Eighty per cent of respondents placed the IFβ-1b scenario as their first choice, and the remaining 20% placed the placebo scenario as the first preference. The most frequent second preference was the placebo scenario (80%), with the IFβ-1b scenario preferred by 16%, followed by the severe scenario (4%). Almost all patients (96%) ranked the severe scenario as the least preferred choice, though two respondents (4%) felt that the IFβ-1b scenario was the worst option.

In terms of EQ-5D states, a large number of respondents ranked 11112 (98%), 11122 (92%) and 21222 (84%) as their first, second and third preferences, respectively. State 22222 was deemed to be the least preferred by 66%, with 21312 ranked as worst by 32%.

The mean health state utility scores produced by the TTO technique are presented in *Table 12*. The utility values elicited for the EQ-5D states ranged from 0.881 for the most preferred state to 0.684 for the least preferred state. The values produced by the University of York MVH group for the same states are given in the final column. The values produced by the respondents in this study are consistently higher than those produced by the general population, with greater differences among the more severe states. The utility scores for the IFβ-1b and placebo scenarios were close to 1, with a difference of 0.014 between them. The severe scenario was valued much lower at 0.359.

As a way of validating this difference of 0.014 between the IFβ-1b and placebo scenarios, different values for the severe scenario were used. The rationale for this is that values for the severe scenario acted as a calibration point in the calculation of utility

TABLE 11 *Ranking of disease-specific scenarios and EQ-5D states*

	Percentage ranked				
	First	**Second**	**Third**	**Fourth**	**Fifth**
Disease-specific scenarios					
IFβ-1b	80	16	4	–	–
Placebo	20	80	0	–	–
Severe	0	4	96	–	–
EQ-5D states					
11112	98	2	0	0	0
11122	2	92	0	6	0
21222	0	6	84	8	2
21312	0	0	16	52	32
22222	0	0	0	34	66

See Appendix 1 for description of scenarios

TABLE 12 *Utility values of disease-specific scenarios and EQ-5D states*

	Mean utility value	**SD**	**Mean MVH value**
Disease-specific scenarios			
IFβ-1b	0.870	0.200	–
Placebo	0.856	0.195	–
Severe	0.359	0.341	–
EQ-5D states			
11112	0.881	0.169	0.829
11122	0.861	0.178	0.722
21222	0.799	0.206	0.553
21312	0.697	0.282	0.536
22222	0.684	0.285	0.500

values for IFβ-1b and placebo scenarios. Clearly therefore, changes in the value of the severe scenario may influence the size of difference between the IFβ-1b and placebo scenarios.

In the TTO interviews, respondents were asked to consider that the severe scenario occurred 'toward the end of a person's life'. Respondents may then have been more willing to trade life years compared with an approach where the scenario occurred over the next 3 years of life. A question posed in this format may have resulted in greater reluctance to lose life expectancy, which would have produced higher utility values. Therefore, the method chosen in this study may have led to values lower than those which reflect true preferences.

In order to gauge the impact on utility values, the value for the severe scenario was set to 0.6 for all respondents, and all values for the IFβ-1b and placebo scenarios were subsequently re-calculated. Mean utility values were calculated as 0.904 for the placebo scenario and 0.917 for the IFβ-1b scenario giving a difference of 0.013, very similar to the original value. This suggests that it is reasonable to conclude that the difference between the IFβ-1b and placebo scenarios is robust to the way in which this question was asked.

Finally, data were collected on the acceptability of the interview. Twelve respondents (24%) thought the tasks asked of them were 'fairly difficult'. However, the vast majority (88%) did not find the interview distressing. The health descriptions were thought realistic by 86% and 82% felt they contained sufficient information.

Chapter 7

Results – direct costs

Average resource use per patient in remission was £529 compared with £2644 per patient in the relapse group. Thus on average, the additional costs associated with a relapse were £2115 per patient. This difference is almost entirely due to inpatient and day case treatment costs and, to a lesser extent, greater receipt of community services. Each group made similar numbers of outpatient visits, and the range of drugs, tests and appliances was comparable.

Tables 13 and *14* show the average total resource cost for the remission (n = 60) and relapse (n = 40) groups for different levels of EDSS score. Complete cost data were not recorded for two of the remission patients, so all of the reported remission costs are based on a sample of 60 people. Higher resource use is associated with higher levels of disability as indicated by the EDSS score, in particular EDSS > 6. However, some people with EDSS below 3 also incurred high costs, due to use of services such as MRI.

In order to test the relationship between cost and disability status, a simple regression analysis was carried out, using cost as the dependent variable, and EDSS score, age and remission/relapse status as independent variables. The results are shown in *Table 15*. They confirm that cost is greater for patients with higher EDSS and having had a recent relapse, but that there is no effect of age.

TABLE 13 *Total direct costs per remission patient (n = 60)*

EDSS	Number of patients	Mean direct costs (£)
0	1	163
1	7	208
1.5	2	811
2	7	426
2.5	3	169
3	7	369
3.5	3	360
4	6	425
4.5	1	989
5	5	784
5.5	2	57
6	8	794
6.5 & 7	3	2296
Not known	5	112
All	**60**	**529**

TABLE 14 *Total direct costs per relapse patient (n = 40)*

EDSS	Number of patients	Mean direct costs (£)
1.5	1	974
2	1	1518
2.5	1	3137
3	3	1631
4	4	2002
5	5	1370
5.5	6	4075
6	6	1490
6.5 & 7	7	5574
Not known	6	1332
All	**40**	**2644**

TABLE 15 *Regression analysis of costs with age, EDSS and remission/relapse status as independent variables*

Dependent variable: log of total cost				
Variable	**Co-efficient**	**Standard error**	**t-value**	**Significance**
Constant	8.140	1.337	6.09	$p < 0.001$
Age	0.009	0.023	0.39	$p = 0.697$
EDSS score	0.333	0.121	2.75	$p < 0.001$
Group*	–1.887	0.443	–4.26	$p < 0.001$
Adjusted $R^2 = 0.320$, $F = 14.79$ (Significance $p < 0.001$)				

* *Group variable: 1 = Remission, 0 = Relapse*

Table 16 shows a breakdown of the type of healthcare resources used. For each element the average cost per patient in each group is compared. Three patients (5%) in the remission group required inpatient admissions within the 6-month study period. These were to specialities other than neurology (orthopaedics, ophthalmology and rehabilitation medicine), involving patients with high EDSS scores (5, 6 and 7, respectively). The average cost per patient was £125.

By definition, all patients in the recent relapse group had at least one admission to hospital either as an in-patient or day case. Twenty-one patients

TABLE 16 *Total direct costs per patient by service and by group*

Service	Mean direct costs per remission patient (£) (n = 60)	Mean direct costs per relapse patient (£) (n = 40)	Difference (£)
Inpatient	125	1212	1087
Day case	20	455	435
Outpatient	128	158	30
Drugs	33	67	34
Procedures	6	66	60
Tests	43	70	27
Appliances	4	13	9
Community	170	603	433
Total	**529**	**2644**	**2115**

(53%) in the recent relapse group received inpatient care, 19 in neurology, one in gynaecology and one in general medicine. The total cost per inpatient admitted ranged from £462 to £10,804 with an average cost of £1212. Median length of stay was 3 days, but longer episodes of care lasting 16, 20 and 73 days were also observed. One patient received both an inpatient (gynaecology) and a day case (neurology) admission.

Four patients (7%) in the remission group were admitted as day cases to neurology or ophthalmology. The average cost per remission patient was £20. Twenty-one patients (53%) in the relapse group received day case admissions, 20 in neurology and one in rehabilitation medicine. The average cost per relapse patient was £455.

A total of 77% of remission patients and 90% of patients in the relapse group made outpatient visits in the previous 6 months. Within the remission group most patients paid one or two visits to neurology, some of whom also made visits to physiotherapy and rehabilitation medicine. Total costs per patient attending as out-patients ranged from £60 to £738 with an average of £128. Patients in the relapse group made relatively more outpatient visits but had a similar distribution across specialties. Outpatient costs per relapse patient ranged from £69 to £1356 with an average of £158.

A total of 57% of remission patients and 5% of relapse patients did not require drug prescriptions within the study period. Total drug cost per patient in both groups was relatively low. The range of total drug cost per patient within each group was broad, however, ranging from £0.78 to £514.48 in the remission group and from £4.20 to £473.20 in

TABLE 17 *Type and frequency of prescribed drugs by group*

Drug	No. of prescriptions (% of all prescriptions)	
	Remission group (n = 60)	Relapse group (n = 40)
Prednisolone	2 (4)	31 (25)
Baclofen	11 (21)	15 (12)
Ditropan	4 (8)	8 (6)
Amitriptyline (sensory symptoms)	4 (8)	3 (2)
Cephalexin	1 (2)	6 (5)
Trimethoprim	1 (2)	5 (4)
Co-codamol	0 (0)	4 (3)
Bromocriptine	3 (6)	1 (1)
Amitriptyline (antidepressant)	2 (4)	2 (2)
Other drugs	25 (47)	49 (40)
Total number of prescriptions	**53**	**124**

the relapse group. The average cost of drugs per relapse patient (£67) was over twice the average cost of drugs per remission patient (£33). *Table 17* shows the most frequently prescribed drugs for both groups. In the remission group a total of 31 different drugs were prescribed, but the total number of prescriptions was low (53). More drugs were prescribed more frequently in the relapse group. A total of 124 prescriptions for 47 different drugs were observed. The most frequently prescribed drugs were all low cost. The highest cost drugs in the study were vitamin B_{12} injections and methylprednisolone; however, only one prescription for each was recorded.

Few procedures were carried out on patients in the remission group. Within the 6-month study period, one lumbar puncture, one lens implant, one therapeutic injection and one cauterisation were observed. Mean procedure cost per remission patient was £6. Mean procedure cost per relapse patient was £66.

Few diagnostic tests were carried out on patients in the remission group, with 80% of patients having no tests at all. The average cost of tests per remission patient was £43. Sixty-five per cent of patients in the relapse group had diagnostic tests with an average cost per patient of £70.

Table 18 shows the most frequently performed diagnostic tests for both groups of patients. Those patients taking part in clinical trials have been excluded as additional testing is used to monitor

TABLE 18 *Type and frequency of diagnostic tests by group (excluding trial patients)*

Test	No. of tests (% of all tests)	
	Remission group (n = 60)	**Relapse group (n = 40)**
Full blood count	1 (3)	17 (12)
Urea and electrolytes	1 (3)	16 (11)
Liver function	2 (6)	12 (8)
Plasma glucose	1 (3)	13 (9)
Urine microbiology	3 (9)	9 (6)
MRI scan	2 (6)	4 (3)
Cholesterol	2 (6)	7 (5)
Erythrocyte sedimentation rate	1 (3)	7 (5)
Thyroid function	2 (6)	6 (4)
Competitive enzyme-linked *Treponema pallidum* immunoassay	0 (0)	6 (4)
Chest X-ray	0 (0)	5 (3)
Cystometrogram	2 (6)	1 (1)
Immunoglobulin μ/ immunoglobulin α	0 (0)	4 (3)
Other tests	18 (51)	36 (25)
Total number of tests	**35**	**143**

their progress. In the remission group a total of 35 tests (27 different tests) were performed. The single, most frequent test was the urine microbiology test. Considerably more diagnostic tests were observed in the relapse group. A total of 143 tests (40 different tests) were observed. The most frequent test was the full blood count. Most of those undergoing tests had at least four, which included a full blood count, urea and electrolytes, a liver function test and a plasma glucose test. Slightly fewer also had the urine microbiology test. Very few high-cost tests were recorded in either group. MRI scan was the most expensive diagnostic test observed in the study.

Five patients (8%) in the remission group were issued with wheelchairs, which generated an average appliance cost of £4. Ten patients (25%) in the relapse group required wheelchairs within the study period, which produced an average cost of appliances per recent relapse patient of £13.

Table 19 shows the frequency with which community services were used. The relapse group received more services than the remission group, particularly for home care and physiotherapy. The mean costs among the entire remission group for home care and physiotherapy were £16 and £19, respectively, whereas for the relapse group the mean costs were £316 and £55, respectively. Total mean community costs per remission patient were £170 compared with £603 per recent relapse patient.

TABLE 19 *Frequency of receipt of community services*

Service	No. of receipts (% of all receipts)	
	Remission group (n = 60)	**Relapse group (n = 40)**
General practitioner	44 (52)	33 (35)
Home care	4 (5)	10 (11)
District nurse	19 (23)	13 (14)
Health visitor	0 (0)	3 (3)
Physiotherapy	7 (8)	14 (15)
Occupational therapy	6 (7)	9 (10)
Speech therapy	0 (0)	4 (4)
Social worker	4 (5)	7 (8)
Total receipts of services	**84**	**93**

Chapter 8

Results – cost-effectiveness and cost–utility analysis

In the results presented here, the cost figures have been rounded to the nearest £100 to minimise the spurious accuracy produced by the division of large numbers by small numbers. The cost-effectiveness calculations detailed in the tables cannot therefore be directly reproduced from the respective figures on cost and effectiveness.

Decision analytic model

Table 20 shows the results of the simple model, which takes into account only the effect on the number of relapses over 5 years. Both cost-effectiveness and cost–utility ratios can be calculated. For each, the figures are presented with and without discounting. The total discounted net costs of IFβ-1b of £43,600 per patient over 5 years are offset by a reduction in relapses by 1.52, giving a cost-effectiveness ratio of £28,700 per relapse avoided. The relapses translate into a gain of 0.054 discounted QALYs, giving a cost–utility ratio of £809,900 per QALY gained. Discounting makes very little difference to these results.

TABLE 20 *Cost and QALY estimates from a 5-year simple model with no effect on progression*

	Standard care	IFβ-1b	Difference
Relapses	5.23	3.71	1.52
Undiscounted costs:			
Relapse	£11,000	£7800	£3200
IFβ-1b	£0	£52,500	£52,500
Total costs	£11,000	£60,300	£49,300
Cost per relapse avoided			**£32,400**
Discounted costs:			
Relapse	£10,000	£7100	£2900
IFβ-1b	£0	£46,600	£46,600
Total cost	£10,000	£53,600	£43,600
Cost per relapse avoided			**£28,700**
QALY gain:			
Undiscounted	–0.204	–0.145	0.059
Discounted	–0.184	–0.130	0.054
Cost per QALY gained:			
Undiscounted			**£831,400**
Costs only discounted			**£736,200**
Costs and QALYs discounted			**£809,900**

The cost figures above have been rounded to the nearest £100 and therefore the cost-effectiveness figures above cannot be directly calculated from the respective figures on costs and effectiveness.

Table 21 shows the results of the 5-year decision analytic model, incorporating changes in progression and using the assumptions presented in *Table 1* (see page 17). Additional costs of IFβ-1b of £43,400 are offset by gains in QALYs of 0.13, indicating a cost–utility ratio of £328,300 per QALY gained. The robustness of this point estimate was tested using a number of one-way sensitivity analyses. *Table 22* shows that different assumptions regarding a range of variables produced no important changes to the size of the cost–utility ratio. The largest changes were observed with respect to the frequency and duration of relapses and their associated utility loss. In addition, more conservative assumptions (i.e. less favourable assumptions from the perspective of the new therapy) produced a higher ratio, which is evidence for the reliability of the model.

The QALY gains estimated using the results from the scenario TTO exercise are considerably lower than those generated by the model, 0.07 rather than 0.13 over 5 years. If this estimate is used, the cost–utility ratio increases to £606,200 per QALY gained, nearly twice the figure generated using the 5-year models.

Table 23 shows the results of the 10-year model, which under base case assumptions produced a cost–utility ratio of £228,300. This model produces similar results to the 5-year model, but as it is slightly more favourable to therapy the 10-year model was analysed in further detail. (However, it is also subject to greater uncertainty; in particular over transition probabilities and therapeutic effects.) Two types of sensitivity analysis were carried out on the 10-year model. First, *Table 23* shows the results from changing a range of variables under favourable ('best case') and unfavourable ('worst case') conditions. The best-case conditions produced an estimate of £74,500 per QALY gained. This incorporated very optimistic

TABLE 21 *Cost and QALY estimates from 5-year decision analytic model (costs discounted)*

EDSS	Standard care		IFβ-1b		Net costs (£)	QALY gains	Costs per QALY gained (£)
	Costs (£)	QALY	Costs (£)	QALY			
3→3	6400	1.50	34,800	2.04			
3→4	2500	0.55	7000	0.40			
3→5	1300	0.23	3600	0.17			
3→6	5000	0.80	13,200	0.60			
Total	**15,200**	**3.08**	**58,600**	**3.21**	**43,400**	**0.13**	**328,300**

The cost figures above have been rounded to the nearest £100 and therefore the cost-effectiveness figures above cannot be directly calculated from the respective figures on costs and effectiveness.

TABLE 22 *Sensitivity analysis of 5-year model*

Parameter	Range of values	Costs per QALY gained (£)
IFβ-1b EDSS transition probabilities		
3→4	0.20–0.10	231,700–366,300
3→5	0.10–0.05	
3→6	0.10–0.25	
Speed of progression: EDSS utilities		
3	0.65–0.79	292,500–380,000
4	0.59–0.74	
5	0.41–0.67	
6	0.38–0.63	
Speed of progression: EDSS costs per year		
3	370–1480	329,200–326,700
4	425–1700	
5	785–3140	
6	795–3180	
Number of relapses with IFβ-1b therapy	2.5–5	223,800–587,600
Length of relapse (weeks)	2–6	431,600–283,200
Utility loss per relapse	0.25–0.75	431,600–264,900
Relapse cost (£)	1000–3000	341,100–318,200
IFβ-1b costs per year (£)	6000–12,000	176,500–379,000
Discount rate	0–10	326,400–328,500

views of the natural history of MS and of the effectiveness of IFβ-1b. The worst case gave an estimate of £604,600 per QALY gained. This was equivalent to the 'no progression' 5-year model.

Second, threshold analysis was carried out to assess the changes in assumptions that would be required in order to make IFβ-1b therapy as cost-effective as current routinely provided therapies. An arbitrary criterion of £50,000 per QALY was adopted and a set of assumptions was generated that would reach the threshold. The assumptions involved changes to three parameters: transition probabilities, speed of progression and severity of relapses. *Table 24*

TABLE 23 *Cost and QALY estimates of 10-year decision analytic model (costs discounted)*

EDSS	Standard care		IFβ-1b		Net costs (£)	QALY gains	Costs per QALY gained (£)
	Costs (£)	QALY	Costs (£)	QALY			
Base case							
3→3	4465	1.21	30,114	2.05			
3→4	5325	1.34	18,167	1.17			
3→5	3205	0.58	9424	0.49			
3→6	16,866	2.19	47,002	1.94			
Total	**29,900**	**5.32**	**104,700**	**5.65**	**74,800**	**0.33**	**228,300**
Best case							
3→3	4500	1.20	41,200	2.83			
3→4	5300	1.32	15,100	1.00			
3→5	3200	0.57	8200	0.48			
3→6	16,900	2.16	37,700	1.92			
Total	**29,900**	**5.25**	**102,200**	**6.23**	**72,300**	**0.98**	**74,500**
Worst case							
3→3	4500	1.21	18,100	1.23			
3→4	5300	1.34	21,200	1.36			
3→5	3200	0.58	11,500	0.60			
3→6	16,800	2.19	54,700	2.26			
Total	**29,900**	**5.32**	**105,400**	**5.45**	**75,500**	**0.13**	**604,600**

Notes: 'Best case' analysis calculates probability of transitions EDSS 3→4, EDSS 3→5, EDSS 3→6 for IFβ-1b using equation $pb = pp - [pp \times 0.5(wp - wb / wp)]$*, where pb is the probability of transition associated with IFβ-1b therapy, pp is the probability of transition associated with standard care, wp and wb are the proportion of standard care and IFβ-1b patients worsening by at least one point on the EDSS instrument respectively. wp = 0.55 and wb = 0.40 using estimates published in reference 1, with pb (EDSS 3→3) =* $1 - \sum pb$*. This gives pb (EDSS 3→3) = 0.41; pb (EDSS 3→4) = 0.15; pb (EDSS 3→5) = 0.08; pb (EDSS 3→6) = 0.36.*

In addition, speed of progression is altered for IFβ-1b, such that (EDSS 3→4) now involves 5 years at EDSS 3 and 5 years at EDSS 4; with (EDSS 3→5) involving 3 years at EDSS 3, 3 years at EDSS 4 and 4 years at EDSS 5; and with (EDSS 3→6) involving 1 year at EDSS 3, 2 years at EDSS 4, 2 years at EDSS 5 and 5 years at EDSS 6.

'Worst case' analysis assumes no difference in the probability of progression between standard care and IFβ-1b, i.e. pb (EDSS 3→3) = 0.18; pb (EDSS 3→4) = 0.21; pb (EDSS 3→5) = 0.11; pb (EDSS 3→6) = 0.50.

compares the assumptions of this threshold case with that of the base case. To highlight the importance of each parameter in generating QALY gains, one-way, two-way and three-way analyses are presented in *Table 25*. This shows that the greatest impact on QALY gains is from improved transition probabilities.

Finally, although the study was carried out specifically to consider the cost-effectiveness of IFβ-1b, the model and data can also be applied to other drug therapies for MS. A 2-year decision analytic model was used to calculate the cost-effectiveness of IFβ-1a and copolymer-1, using data published from trials of their effectiveness.[74,75] (Five-year trial data are not yet available.) This model was also applied to IFβ-1b for comparison. *Table 26* shows the assumptions used. The costs of the drugs are, or in the case of copolymer-1 are likely to be, very similar to each other. Therefore, in order to avoid inappropriate calculations based on different pricing strategies, the models all assume that the price is the same as for IFβ-1b. Because of this, and equally importantly because the three trials and the data published from them have a number of important differences, the results should not be taken as a direct comparison of the three therapies in terms of their cost-effectiveness. However, the results, shown in *Table 27*, demonstrate that the cost-effectiveness of these other drugs is likely to be very similar to that of IFβ-1b if their price is similar.

TABLE 24 *Assumptions used in threshold analysis compared with base case*

Parameter	Threshold		Base case	
	Standard care	IFβ-1b	Standard care	IFβ-1b
EDSS transition probabilities				
3→3	0.18	0.55	0.18	0.30
3→4	0.21	0.20	0.21	0.18
3→5	0.11	0.10	0.11	0.09
3→6	0.50	0.15	0.50	0.43
Speed of progression: mean EDSS transition utilities per year				
3→3	0.71	0.71	0.71	0.71
3→4	0.67	0.68	0.67	0.67
3→5	0.57	0.62	0.57	0.57
3→6	0.48	0.55	0.48	0.48
Speed of progression: mean discounted EDSS transition costs per year (£)				
3→3	577	577	577	577
3→4	632	614	632	632
3→5	1010	818	1010	1010
3→6	1470	1017	1470	1470
Severity of relapse: utility loss per relapse	0.5	0.2	0.5	0.5

TABLE 25 *Effect of parameters used in threshold analysis on QALY gains and costs per QALY gain*

Type of sensitivity analysis	Parameters			QALY gains	Costs per QALY gained (£)
	Transition probabilities	Speed of progression	Severity of relapse		
Base case				0.33	228,300
One-way			✔	0.48	130,200
		✔		0.67	94,600
	✔			0.91	70,400
Two-way		✔	✔	0.87	74,000
	✔	✔		1.10	58,800
	✔		✔	1.11	58,300
Three-way	✔	✔	✔	1.30	50,000

TABLE 26 *Assumptions of a 2-year model for IFβ-1b, IFβ-1a and copolymer-1*

Parameter	Model assumptions					
	Standard care	IFβ-1b	Standard care	IFβ–1a	Standard care	Copolymer-1
EDSS transition probabilities						
3→3	0.71	0.78	0.69	0.82	0.71	0.79
3→4	0.10*	0.13*	0.11	0.11	0.10*	0.13*
3→5	0.19*	0.09*	0.20	0.07	0.19*	0.08*
EDSS utilities						
3	0.71	0.71	0.71	0.71	0.71	0.71
4	0.66	0.66	0.66	0.66	0.66	0.66
5	0.52	0.52	0.52	0.52	0.52	0.52
EDSS costs per year (£)						
3	740	740	740	740	740	740
4	850	850	850	850	850	850
5	1570	1570	1570	1570	1570	1570
Relapse rate per year	1.27	0.84	0.90	0.61	0.84	0.59
Relapse length (months)	1	1	1	1	1	1
Utility loss per relapse	0.5	0.5	0.5	0.5	0.5	0.5
Relapse cost (£)	2115	2115	2115	2115	2115	2115
Drug costs per year (£)	0	10,500	0	10,500†	0	10,500†

* *Transition probabilities EDSS 3→4, EDSS 3→5 for IFβ-1a and copolymer-1 derived using IFβ-1b data.*
† *Drug costs per year for IFβ-1a and copolymer-1 are assumed equal to those of IFβ-1b.*

TABLE 27 *Costs and QALY estimates of a 2-year decision analytic model applied to IFβ-1b, IFβ-1a and copolymer-1*

EDSS	Standard care		Therapy		Net costs (£)	QALY gains	Costs per QALY gained (£)
	Costs (£)	QALY	Costs (£)	QALY			
IFβ-1b							
3→3	4865	0.93	20,306	1.05			
3→4	696	0.13	3399	0.17			
3→5	1481	0.20	2427	0.10			
Total	**7000**	**1.26**	**26,100**	**1.32**	**19,100**	**0.06**	**327,300**
IFβ-1a							
3→3	3648	0.93	20,549	1.12			
3→4	594	0.14	2769	0.15			
3→5	1245	0.22	1820	0.08			
Total	**5500**	**1.29**	**25,100**	**1.35**	**19,600**	**0.06**	**354,900**
Copolymer-1							
3→3	3574	0.96	19,731	1.08			
3→4	514	0.13	3261	0.17			
3→5	1135	0.21	2073	0.09			
Total	**5200**	**1.30**	**25,100**	**1.35**	**19,900**	**0.05**	**433,900**

Chapter 9
Discussion

Quality of life

The instruments used in this study demonstrated that it is possible to quantify in a number of ways, the impact of MS on patients' QOL. The overall effect of MS *per se* is measurable, as demonstrated by the comparison of relapse and remission groups, and of people in the general population with or without long-standing illness. The effects on QOL of having a relapse are also measurable. This suggests that there is no real justification for the exclusion of QOL measures in evaluations of therapies for MS.

Both MSQOL-54 and EQ-5D detected that the relapse group had poorer QOL than the remission group, both when the relapse was at its worst and 'today'. This suggests that the effects of relapse may continue for several months. However, the effects of relapses are complex and require close examination, as patients may have experienced relapses at different times over a 6-month period, and therefore, some may have fully recovered from relapse and entered a new relapsing-progressive phase. The use of summary measures of effectiveness such as relapse rates or EDSS progression may therefore mask important QOL implications.

As the diary data had to be collected prospectively, only patients in remission were requested to provide information. Analyses conducted so far suggest only a small number of patients experienced symptom-free days over a 6-week period. In addition, patients with poorer EDSS status reported more symptoms than those with lower scores. The diary therefore seemed to be picking up an 'iceberg of ill-health', including problems not severe enough to require contact with health professionals. This suggests that changes in the health status of people with MS may have been inadequately measured by the outcome measures normally used in clinical trials. However, further analysis in this context is required to assess whether the diary method itself is valuable in providing information on fluctuations in health status at an individual level.

Differences on all three measures (MSQOL-54, EQ-5D and diaries) in relation to EDSS scores were as expected, with higher EDSS scores associated with poorer QOL profiles. QOL measures are therefore consistent with this clinical measure, which is important as EDSS is so widely used and accepted by clinicians. However, QOL assessment provided important additional information about dimensions of QOL clearly affected by having a relapse and by disease progression.

Finally, patients' valuation of EQ-5D health states were consistently higher compared with those of the general population, suggesting they value poorer states less unfavourably than the general population, perhaps due to greater experience of ill-health. In addition, utility gains estimated directly from patients' valuation of scenarios were very different from those estimated via EQ-5D and the York tariff. This will have implications in considering the issue of which values should be used in an economic evaluation and how they should be measured.

Direct costs

There was a substantial difference in the level of resource use and costs between the remission and relapse groups. Therefore therapy which reduces the number of relapses will have a favourable impact on costs. There was a positive relationship between costs and EDSS, with higher EDSS associated with greater resource use in remission. Therefore, therapy which delays progression will also have a favourable impact on costs. However, both of these cost reductions are small compared with the cost of IFβ-1b therapy, resulting in a substantial net increase in costs associated with its use.

Cost-effectiveness and cost–utility analyses

Current trial-based evidence of the impact of IFβ-1b on people with RRMS is restricted to a statistically significant reduction in the number of relapses. Considering this evidence alone leads to a cost-effectiveness ratio of £28,700 per relapse avoided by the use of IFβ-1b. This translates into a 'worst-case' cost–utility ratio of £809,900 per QALY gained. Allowing for effects on progression over 5 years, which were, however, not found to be statistically significant, gives a cost–utility ratio of £328,300 per

QALY gained. Extrapolating these gains beyond the duration of the trials to 10 years gives a cost–utility ratio of £228,300 per QALY gained.

Sensitivity analysis carried out on these figures suggested that the cost–utility estimates are robust to changes in assumptions, and also provided evidence of the validity of the model. The range of estimates assuming an effect on progression was between £176,500 and £587,600 per QALY gained for the 5-year model, and the most optimistic estimate from the 10-year model, incorporating some extreme assumptions about natural history and impact of therapy, was £74,500 per QALY gained.

The CCOHTA report also presented cost-effectiveness ratios based on the cost per relapse avoided and the cost of avoiding disability.[2] The study on which this report was based used a number of different methods and assumptions, including a different source of natural history data. Unfortunately, the cost of avoiding disability cannot be directly compared, because the study did not translate EDSS changes into QALYs, rather quoting costs per normalised EDSS disability year avoided. However, the cost per relapse figures are directly comparable. They quoted a cost per relapse avoided in the range $48,000 to $67,000, which converts to £20,000 to £28,000 at current exchange rates. This is slightly lower than the figures reported here. However, the reported cost of IFβ-1b in Canada, from where the calculations were made, is lower than that in the UK; using the same exchange rate, the annual cost is £7000 rather than £10,000. Using this lower cost figure in our model gives a cost per relapse avoided of £20,000, exactly the same as the lower Canadian estimate.

The above findings, together with the findings from the QOL surveys, suggest that IFβ-1b produces important occasional short-term gains in QOL to people with RRMS. However, because of the infrequency of these gains they translate into only small gains in QALYs overall. In addition, because many people would not benefit from the longer-term gains that optimistic estimates suggest would come from a reduced probability of progression, the aggregate QALY gains will also be small. As all of these benefits are only achieved with a large additional cost compared with standard clinical management, IFβ-1b has a high cost per QALY gained. Preliminary evidence suggests that other drug therapies currently being tested may have similar levels of cost-effectiveness.

Whether it is appropriate or not to call these costs per QALY gained high depends on the definition of 'high'. The criteria adopted for judging such matters are necessarily based on value judgements and are context dependent. Although criteria for cost-effectiveness have been advocated,[76] there are currently no defined standards applicable to the NHS. As this study is of cost-effectiveness rather than cost–benefit, 'high' therefore means relative to many other known interventions and not to an absolute value.

Limitations of the study

The main limitation of this study is that it was not conducted as part of a clinical trial. Reliance on existing trial data meant that a number of assumptions had to be made to convert the clinical evidence into an economic evaluation. These assumptions were evidence-based, plausible and robust to testing by sensitivity analysis, but undoubtedly better information could be obtained, producing more precise estimates.

A second limitation is that some key instruments for the collection of new data were not available and had to be developed in a short space of time, in particular the QOL instruments. Although these instruments performed well, they require further refinement. Data collection had similar time constraints. This meant that a number of data could not be collected (e.g. cost data for people with EDSS > 7), that some data had to be collected retrospectively (e.g. case notes for EDSS and resource use), and that the data collection design could not be planned in accordance with proper statistical principles to ensure power and significance.

In addition, the results are only applicable to RRMS and cannot be generalised to other categories of MS. However, the model itself does take account of progression in line with the published clinical trial data on which it is based.[1,56]

Finally, the analysis exclusively focused on IFβ-1b, with only a demonstration of the applicability of the model to IFβ-1a and copolymer-1. It was not possible to compare our results with the costs and benefits of other interventions used in MS, such as treatments used during relapses or provision of a dedicated MS service.

Conclusions

Implications for policy

The implications of the cost-effectiveness results rely on value judgements, which the authors of this

report have no grounds for making. Prescribing and policy decisions need to be taken on the basis of judgement at a number of different levels, and the results of the study may be helpful to those making the decisions, for example, MS patients and their clinicians, and health authorities.

At the level of individual decision making by patients and clinicians, it may be important to consider the true extent of the gains in QOL produced by IFβ-1b and to consider these in the light of the alternatives for improving the QOL of people with MS. The cost-effectiveness calculations given here (which do not consider such alternatives) may be less important than the absolute levels of benefits and costs, which might be used to inform an individual cost–benefit calculation.

Health authorities ought to consider alternative ways in which funds that might be spent on IFβ-1b could be used to improve the QOL of people with MS. However, they also have to consider whether the extra investment required is worthwhile in terms of producing reductions in relapses and QALY gains, compared with the gains that health care produces for people with other conditions. It will be important to bear in mind both efficiency, as indicated by the cost–utility figures, but also equity. The latter might take into account the current low level of spending on people with MS and the difficulty of producing QALY gains by other means.

Implications for research

This study has implications for future studies of MS and for the impact of new therapies on this disease. In particular, it attaches importance to placing QOL measurement at the heart of outcomes measurement, and the need to link trial data with natural history and cost data. These aspects would probably have a substantial impact on trial design.

Acknowledgements

The project described in this report was commissioned by the NHS Health Technology Assessment programme, project number 95/01/2. It was conducted by the following project team at the University of Newcastle-upon-Tyne:

Dr David Bates, Senior Lecturer and Consultant Neurologist, Department of Clinical Neurology; Ms Julie Doughty, Junior Research Associate, Centre for Health Services Research; Sister Joanna Forsyth, Staff Nurse, Department of Clinical Neurology; Mrs Sylvia Hudson, Project Secretary, Centre for Health Services Research; Dr Ann Jacoby, Principal Research Associate, Centre for Health Services Research; Mr Paul McNamee, Research Associate, Centre for Health Services Research; Mr Paul Miller, Research Associate, Department of Epidemiology & Public Health; Dr David Parkin, Senior Lecturer in Health Economics, Department of Epidemiology & Public Health; Dr Simon Thomas, Consultant in General Medicine and Clinical Pharmacology, Wolfson Unit of Clinical Pharmacology.

We are grateful to the people who gave their time to take part in our surveys, to Paul Dolan and Karen Gerard for help with the valuation exercise, to Pippa Anderson for access to psychometric analysis software and to referees of this report for their helpful comments.

References

1. The IFNB Multiple Sclerosis Study Group and University of British Columbia MS/MRI Analysis Group. Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomised controlled trial. *Neurology* 1995;**45**:1277–85.
2. Canadian Coordinating Office for Health Technology Assessment (1995). Interferon beta-1b and multiple sclerosis. CCOHTA Technology Overview, Ottowa, December 1996.
3. Feinstein AR, Josephy BR, Wells CK. Scientific and clinical problems in indexes of functional disability. *Ann Intern Med* 1986;**105**:413–20.
4. McCartney CF, Larson DB. Quality of life in patients with gynaecological cancer. *Cancer* 1987;**60**:2129–36.
5. Spilker B. Quality of life and pharmacoeconomics in clinical trials. 2nd edition. Philadelphia: Lippincott-Raven, 1996.
6. Read JL. The new era of quality of life assessment. In: Walker SR, Rosser RM, editors. Quality of life assessment: key issues in the 1990s. Lancaster: Kluwer Academic Publishers, 1993.
7. Schipper H, Clinch J, McMurray A, Levitt M. Measuring the quality of life of cancer patients: the Functional Living Index – Cancer. Development and validation. *J Clin Oncol* 1984;**2**:472–83.
8. Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR. Quality of life assessment: can we keep it simple? *J Royal Stat Soc* 1992;**155**(3):353–93.
9. Fletcher S, Gore S, Jones D, Fitzpatrick R, Spiegelhalter D, Cox D. Quality of life measures in health care. II: Design, analysis and interpretation. *BMJ* 1992;**305**:1145–8.
10. Sprangers MAG, Aronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. *J Clin Epidemiol* 1992;**45**:743–60.
11. Guyatt G, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;**118**:622–9.
12. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;**27**:S217–S232.
13. Bowling A. Measuring disease. Buckingham: Open University Press, 1995.
14. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D. Quality of life measures in health care. I: Applications and issues in assessment. *BMJ* 1992;**305**:1074–7.
15. Fallowfield L. The quality of life: the missing measurement in health care. London: Souvenir Press, 1990.
16. O'Boyle CA, McGee H, Hickey A, O'Malley K, Joyce CRB. Individual quality of life in patients undergoing hip replacement. *Lancet* 1992;**339**:1088–91.
17. Ruta DA, Garratt AM, Leng M, Russell IT, Macdonald LM. A new approach to the measurement of quality of life: the Patient-Generated Index. *Med Care* 1994;**32**:1109–26.
18. Kendrick AM, Trimble MR. Repertory grid in the assessment of quality of life in patients with epilepsy: the quality of life assessment schedule. In: Trimble M, Dodson W, editors. Epilepsy and quality of life. New York: Raven Press Ltd, 1994.
19. Fraser SCA, Ramirez AJ, Ebbs SR, Fallowfield LJ, Dobbs HJ, Richards MA, *et al.* A daily diary for QOL measurement in advanced breast cancer trials. *Br J Cancer* 1993;**67**:340–6.
20. Fayers P. MRC quality of life studies using a daily diary card – practical lessons learned from cancer trials. *Qual Life Res* 1995;**4**:343–52.
21. Hyland ME, Crocker GR. Validation of an asthma quality of life diary in a clinical trial. *Thorax* 1995;**50**:724–30.
22. Verbrugge LM. Health diaries. *Med Care* 1980;**18**:73–95.
23. Gotay CG. Patient-reported assessments versus performance-based tests. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd edition. Philadelphia: Lippincott-Raven, 1996.
24. Corti L. Using diaries in social research. Social Research Update. Issue 2, March 1993. Guildford: University of Surrey, 1993.
25. Davis MZ. Living with multiple sclerosis – a social-psychological analysis. Springfield, Ill: Charles C Thomas, 1973.
26. Schienberg LC. Studies in the long-term psychosocial problems of patients with multiple sclerosis. *Int J Rehab Res* 1979;**2**:540–2.
27. Stewart DC, Sullivan TJ. Illness behaviour and the sick role: the case of multiple sclerosis. *Social Sci Med* 1982;**16**:1397–404.

28. Burnfield A. Multiple sclerosis: a personal exploration. London: Souvenir Press, 1985.

29. Harper AC, Harper DA, Chambers LW, Cino PM, Singer J. An epidemiological description of physical, social and psychological problems in multiple sclerosis. *J Chron Dis* 1986;**39**:305–10.

30. Robinson I. Multiple sclerosis. London: Routledge, 1988.

31. Rudick RA, Miller D, Clough JD, Gragg LA, Farmer RG. Quality of life in multiple sclerosis: comparison with inflammatory bowel disease and rheumatoid arthritis. *Arch Neurol* 1992;**49**:1237–42.

32. Murray TJ. The psychosocial aspects of multiple sclerosis. *Neurolog Clinic* 1995;**13**:197–223.

33. International Federation of Multiple Sclerosis Societies. Minimal record of disability for multiple sclerosis. New York: National Multiple Sclerosis Society, 1985.

34. Noseworthy JH, Vander voort MK, Wong CJ, Ebers GC. Interrater reliability with the Expanded Disability Status Scale (EDSS) and Functional Systems (FS) in a multiple sclerosis clinical trial. *Neurology* 1990;**40**:971–5.

35. Willoughby EW, Paty DW. Scales for rating impairment in multiple sclerosis: a critique. *Neurology* 1988;**38**:1793–8.

36. Whitaker JN, McFarland HF, Rudge P, Reingold SC. Outcomes assessment in multiple sclerosis trials: a critical analysis. *Multiple Sclerosis* 1995;**1**:37–47.

37. Kurtzke JF. Rating neurological impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;**33**:1444–52.

38. Amato MP, Fratiglioni L, Groppi C, Syracusa G, Amaducci L. Interrater reliability in assessing functional systems and disability on the Kurtzke scale in multiple sclerosis. *Arch Neurol* 1988;**45**:746–8.

39. Amato MP, Hours M, Bartolozzi ML, Cortinovis P, Confavreux C. European cross-national standardisation in the use of the EDMUS database system for multiple sclerosis. *J Neurol* 1996;**243**(suppl 2):S82.

40. Verdier-Taillefer MH, Roullet E, Cesaro P, Alperovitch A. Validation of self-reported neurological disability in multiple sclerosis. *Int J Epidemiol* 1994;**23**:148–54.

41. Sipe JC, Knobler RL, Braheny SL, Rice JP, Panitch HF, Oldstone MB. A neurological rating scale (NRS) for multiple sclerosis. *Neurology* 1984;**34**:1368–72.

42. Mumford CJ, Compston A. Problems with rating scales for multiple sclerosis: a novel approach – the CAMBS score. *J Neurol* 1993;**240**:209–15.

43. Sharrack B, Hughes RAC, Soudain S. Guy's Neurological Disability Scale. *J Neurol* 1995;**243**(suppl 2):S32.

44. Cella DF, Dineen K, Arnason B, Reder A, Webster KA, Karabatsos G, *et al.* Validation of the Functional Assessment of Multiple Sclerosis QOL Instrument. *Neurology* 1996;**47**:129–39.

45. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, *et al.* The Functional Assessment of Cancer Therapy (FACT) scale: development and validation of the general measure. *J Clin Oncol* 1993;**11**:570–9.

46. Ford HL, Tennant A, Johnson MH. The Leeds MSQoL scale: a disease-specific measure of quality of life in multiple sclerosis. Personal communication.

47. Ware JE, Sherbourne CD. A 36-item short-form health survey (SF-36) I. Conceptual framework and item selection. *Med Care* 1992;**30**:473–83.

48. Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. *Qual Life Res* 1995;**4**:187–206.

49. Smith K, Wright K. Informal care and economic appraisal: a discussion of possible methodological approaches. *Health Economics* 1994;**3**:137–48.

50. O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Med Care* 1994;**32**(2):150–63.

51. Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomised control trials. *Med Care* 1992;**30**(3):231–9.

52. Bourdette DN, Prochazka AV, Mitchell W, Licari P, Burks J, and VA Multiple Sclerosis Rehabilitation Study Group. Health care costs of veterans with multiple sclerosis: implications for the rehabilitation of MS. *Arch Phys Med Rehab* 1993;**74**:26–31.

53. Wolfson C, Confavreux C. Improvements to a simple Markov model of the natural history of multiple sclerosis. i. Short-term prognosis. *Neuroepidemiology* 1987;**6**:101–15.

54. Confavreux C, Wolfson C. Mathematical models and individualized outcome estimates in multiple sclerosis. *Biomed Pharmacother* 1989;**43**:675–80.

55. Sox HC, Blatt MA, Higgins MC, Marton KI. Medical decision making. Boston: Butterworths, 1988.

56. The IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis: I clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 1993;**43**:655–61.

57. Williams A. The role of the EUROQOL instrument in QALY calculations. Discussion Paper 130. University of York: Centre for Health Economics, 1995.

58. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey. University of York: Centre for Health Economics, 1995.

59. Robinson I. Analysing the structure of 23 clinical trials in multiple sclerosis. *Neuroepidemiology* 1987;**6**:46–76.

60. Bullinger M. Indices versus profiles – advantages and disadvantages. In: Walker SR, Rosser RM, editors. Quality of life assessment: key issues in the 1990s. Lancaster: Kluwer Academic Publishers, 1993.

61. Williams A. Health economics, priority setting and medical ethics: implications for multiple sclerosis. Discussion Paper 71. University of York: Centre for Health Economics, 1990.

62. Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford Medical Publications, 1987.

63. Weinshenker BG, Bass B, Rice GPA, Noseworthy J, Carriere W, Baskerville J, *et al.* The natural history of multiple sclerosis: a geographically based study. i. Clinical course and disability. *Brain* 1989;**112**:133–46.

64. Weinshenker BG, Bass B, Rice GPA, Noseworthy J, Carriere W, Baskerville J, *et al.* The natural history of multiple sclerosis: a geographically based study. ii. Predictive value of the early clinical course. *Brain* 1989;**112**:1419–28.

65. Weinshenker BG, Rice GPA, Noseworthy JH, Carriere W, Baskerville J, Ebers GC. The natural history of multiple sclerosis: a geographically based study. iii. Multivariate analysis of predictive factors and models of outcome. *Brain* 1991;**114**:1045–56.

66. Weinshenker BG, Rice GPA, Noseworthy JH, Carriere W, Baskerville J, Ebers GC. The natural history of multiple sclerosis: a geographically based study. iv. Applications to planning and interpretations of clinical therapeutic trials. *Brain* 1991;**114**:1057–67.

67. Chartered Institute of Public Finance and Accountancy. Health service financial database. London: Health Service Management Association, 1996.

68. British Medical Association, Royal Pharmacological Society of Great Britain. British National Formulary. London: British Medical Association, Royal Pharmacological Society of Great Britain, 1996.

69. Netten A, Dennett J. Unit costs of health and social care. University of Kent at Canterbury: Personal Social Services Research Unit, 1996.

70. Robinson I, Lawson A, Bakes C. Questioning the facts: the sex ratio in multiple sclerosis. Brunel-ARMS Research Unit, Department of Human sciences. Brunel: University of West London, 1983.

71. Ware JE, Harris WJ, Gandek B, Rogers BW, Reese PR. MAP-R for Windows User's Manual (Draft). Boston: The Health Institute, New England Medical Centre, 1997.

72. Jenkinson C, Coulter A, Wright L. Short-form 36 (SF-36) health survey questionnaire: normative data for adults of working age. *BMJ* 1993;**306**:1437–40.

73. Williams A. The measurement and valuation of health: a chronicle. University of York: Centre for Health Economics, 1995.

74. Jacobs LD, Cookfair DL, Ruddick RA, Herndon RM, Richert JR, Salazar AM, *et al.* Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Ann Neurol* 1996;**39**(3):285–93.

75. Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multi centre, double-blinded, placebo-controlled trial. *Neurology* 1995;**45**:1268–76.

76. Laupacis A, Feeny D, Detsky AS, Tugwell PX. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Can Med Assoc J* 1992;**146**(4):473–81.

Appendix 1

Utility assessment scenarios

Multiple sclerosis scenarios

Good health

- Leading an active life and finding work and other interests/activities rewarding
- Having good relations with family and friends
- Leading a healthy lifestyle in terms of diet and leisure activities
- Having a positive approach to life and rarely feeling anxious or depressed
- Having no health problems which cause pain or discomfort
- Welcoming new challenges and feeling optimistic about the future in both work and personal life

Health description 1 (Placebo scenario)

A person has had multiple sclerosis for 5 years. They are currently in remission (that is, not having a relapse) and have:

- no problems walking about
- no problems in self-care (washing or dressing)
- some problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- little or no pain/discomfort
- occasional bouts of anxiety or depression.

Over the next **3** years they will have **three** relapses. These relapses involve the development of new symptoms or the worsening of existing symptoms, either of which last for longer than 24 hours. There is a **25% chance** that the relapses are severe enough to require treatment in hospital.

Towards the end of the 3 years, after the three relapses have occurred, there is a **30% chance** that health while in remission will have deteriorated to an extent that they have:

- greater problems with performing usual activities
- greater (though moderate) pain or discomfort.

Health description 2 (IFb-1b scenario)

A person has had multiple sclerosis for 5 years. They are currently in remission (that is, not having a relapse) and have:

- no problems walking about
- no problems in self-care (washing or dressing)
- some problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- little or no pain/discomfort
- occasional bouts of anxiety or depression.

Over the next **3** years they will have **two** relapses. These relapses involve the development of new symptoms or the worsening of existing symptoms, either of which last for longer than 24 hours. There is a **20% chance** that the relapses are severe enough to require treatment in hospital.

Over the next 3 years they will have weekly intramuscular injections. During the first year of this treatment they may experience mild, occasional flu-like symptoms, muscle aches and chills.

Towards the end of the 3 years, after the two relapses have occurred, there is a **20% chance** that health while in remission will have deteriorated to an extent that they have:

- greater problems with performing usual activities
- greater (though moderate) pain or discomfort.

Health description 3 (Severe scenario)

A person has had multiple sclerosis for 5 years. They are currently in remission (that is, not having a relapse) and have:

- difficulty walking short distances (e.g. 100 yards)
- some problems in self-care (washing or dressing)
- a total inability to perform all usual activities (e.g. work, study, housework, family or leisure activities)
- moderate pain/discomfort
- extreme anxiety or depression.

Over the next **3** years they will have **more than three** relapses. These relapses involve the development of new symptoms or the worsening of existing symptoms, either of which last longer than 24 hours. Most of these relapses are severe enough to require treatment in hospital.

Towards the end of the 3 years, after the relapses have occurred, health will have deteriorated to such an extent that they are confined to a wheelchair or bed on most days, and have:

- greater problems with performing usual activities
- greater (though moderate) pain or discomfort.

EQ-5D health states

Health state 1

- No problems walking about
- No problems with self-care (washing or dressing self)
- No problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- No pain or discomfort
- Moderately anxious or depressed

Health state 2

- No problems walking about
- No problems with self-care (washing or dressing self)
- No problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- Moderate pain or discomfort
- Moderately anxious or depressed

Health state 3

- Some problems walking about
- No problems with self-care (washing or dressing self)
- Some problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- Moderate pain or discomfort
- Moderately anxious or depressed

Health state 4

- Some problems walking about
- No problems with self-care (washing or dressing self)
- Unable to perform usual activities (e.g. work, study, housework, family or leisure activities)
- No pain or discomfort
- Moderately anxious or depressed

Health state 5

- Some problems walking about
- Some problems with self-care (washing or dressing self)
- Some problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- Moderate pain or discomfort
- Moderately anxious or depressed

Appendix 2

Unit costs

	Cost per day (£)
Inpatient costs*	
Neurology	179
General medicine	154
Rehabilitation medicine	148
Gynaecology	266
Orthopaedic	221
Ophthalmology	447
Urology	219
***Day case costs**†*	
Neurology	101
Rehabilitation medicine	300
Ophthalmology	150

* *Based on the CIPFA database.*[67]
† *Based on CIPFA unit cost estimates multiplied by frequency of visit.*

	Cost per visit (£)
***Outpatient costs**†*	
Neurology	69
General medicine	61
Rehabilitation medicine	300
Gynaecology	53
Orthopaedic	51
Ophthalmology	38
Urology	60
Ear, nose and throat	53
Surgery	49
Endocrinology	61

† *Based on CIPFA unit cost estimates multiplied by frequency of visit.*

	Cost per procedure (£)
***Procedure costs**‡*	
Methylprednisolone injection	10
Diagnostic spinal puncture	71

‡ *From local Trust providers.*

	Cost per appliance (£)
***Appliance costs**¶*	
Manual wheelchair	54
Electric wheelchair	164

¶ *Calculated from PSSRU unit costs.*[69]

	Cost per test (£)
***Test costs**‡*	
Liver function	8
Urea and electrolytes	9
Plasma glucose	5
Erythrocyte sedimentation rate	1
Full blood count	3
Oligoclonal bands	25
Auto antibody screening	12
Protein/glucose	5
Cell count and type	3
IgD index	13
Computerised tomography scan	174
MRI scan	250
Ambulatory urodynamic studies	147
Cytometrogram	67
Urethral sphincter electromyography	67
Urine/microbiology	12
Renal ultrasound	22
Thyroid function	9
Electrocardiogram	9
Antigliadin antibodies	5
Bone density scan	207
Cholesterol	4
Ph serum	9
Immunoglobulin μ/ immunoglobulin α protein electrophoresis	14
Paraprotein	5
Ionised calcium	3
Immunofixation	40
Coagulation screen	10
Competitive enzyme-linked *Treponema pallidum* immunoassay	10
Venereal Diseases Research Laboratory	9
Prothrombin time	3
Serum prolactin	5
Plasma B12/folate/red cell folate	6
Mammogram	31
Arthroscopy	85
X-ray hand	13
X-ray chest	20
X-ray shoulder	17
X-ray abdomen	17
X-ray cervical & lumbar	36
Ultrasound breast	25
Aspiration cytology	17

‡ *From local Trust providers.*

	Cost per mg or ml (£)
Drug costs **	
Amantadine/Symmetrel®	0.0019
Amitriptyline	0.0004
Amoxycillin	0.0001
Arthrotec®	0.0050
Aspirin	0.0001
Azathioprine	0.0036
Baclofen	0.0060
Becloforte®	0.1155
Bendrofluazide	0.0040
Bicillin®	0.0075
Bricanyl Inhaler®	0.0133
Bromocriptine (Parlodel®)	0.0720
Buscopan®	0.0080
Canusal® (heparin)	0.0029
Carbamazepine	0.0003
Cefuroxime (Zinnat)	0.0027
Cephalexin	0.0005
Chloramphenicol	0.1310
Cimetidine	0.0003
Ciprofloxacin/Ciproxin®	0.0027
Clexane® (heparin)	0.1183
Co-amilofruse (Frumil)	0.0440
Co-codamol®	0.0002
Co-proxamol®	0.0001
Codeine phosphate	0.0013
Crystapen®	0.0007
Cystrin®/oxybutynin	0.0540
Dantrium® intravenous	1.1260
Dantrolene sodium/Dantrium®	0.0068
Diazepam®	0.0010
Dihydrocodeine®	0.0010
Ditropan®/oxybutynin	0.0560
Dothiepin	0.0016
Flucloxacillin (Floxapen®)	0.0003
Fluoxetine/Prozac®	0.0345
Fragmin® (heparin)	0.0002
Gamanil®	0.0026
Glycerol Suppositories®	0.0001
Hormone replacement therapy (all brands)	0.0450
Hypromellose	0.0850
Imipramine	0.0010
Ketoprofen®	0.0014
Lactulose	0.0053
Lentizol®	0.0018
Lignocaine hydrochloride	0.0001
Lioresal®	0.0110
Manevac granules®	0.0001
Maxepa®	0.0786
Methylprednisolone/Depo-Medrone®	2.7300
Metronidazole	0.0001

** *Calculated using BNF.*[68]

	Cost per mg or ml (£)
***Drug costs** continued*	
Monoparin® (heparin)	0.0014
Multiparin® (heparin)	0.0005
Naprosyn®	0.0005
Nifedipine	0.0120
Nitrofurantoin	0.0019
Omeprazole/Losec®	0.0635
Paroxetine/Seroxat®	0.0347
Phenylephrine eye drops	0.2660
Pred Forte®	0.3190
Prednisolone	0.0100
Predsol®	0.0200
Primolut N®	0.0140
Propantheline/Pro-Banthine®	0.0033
Propranolol	0.0003
Pulmicort Turbohaler®	0.0925
Ranitidine	0.0030
Senna tablets	0.0013
Senokot®	0.0001
Sodium valproate	0.0003
Tegretol®	0.0003
Tegretol Retard®	0.0004
Temazepam	0.0030
Thyroxine	0.0256
Tofranil®	0.0016
Trimethoprim	0.0002
Tryptizol®	0.0004
Uniparin Forte® (heparin)	0.0002
Ursodeoxycholic acid	0.0021
Ventolin®	0.0115
Ventolin Accuhaler®	0.0883
Ventolin Rotacaps®	0.2000
Vitamin B_{12} (injected)	1.3600

** *Calculated using BNF.*[68]

	Cost per visit (£)
Community-based costs ¶	
General practitioner	47
Home care	6
District nurse	22
Health visitor	32
Social worker	15
Occupational therapy	23
Speech therapy	27
Physiotherapy	19

¶ *Calculated from PSSRU unit costs.*[69]

Appendix 3

Psychometric analysis of MSQOL-54

Although its authors provide a considerable amount of information about the psychometric properties of MSQOL-54,[48] this measure was derived from a sample of patients in the USA. We therefore undertook further psychometric analysis using the multitrait scaling analysis package, MAP-R for Windows, recently developed by Ware and co-workers.[71] For any data set, MAP-R computes percentages of missing and computable data, and data quality as measured by item internal consistency and scaling success. The analysis was run on all patients, together and on the remission and relapse groups separately. Week 1 questionnaire data were used for the remission patients.

Data completeness

The number and percentage of patients missing each of the 50 MSQOL-54 items are shown in *Table 28*. Missing value rates ranged from 2.0 (BP3, SF3, QOL1, TRANS) to 8.8 for all three items in the role emotional scale (RE1, RE2, RE3). Data were complete (no missing or out-of-range responses) for only 71 patients (70%). However, MAP-R permits computation of scale scores whenever responses are provided for at least 50% of the items in that scale. In such cases, a subject-specific value is imputed which is the mean value across completed items in the scale after re-coding. The authors consider this approach satisfactory for testing data quality of an established questionnaire. Within this boundary, scales were computable for 88 patients (86%).

Data quality

One measure of the quality of data obtained is item internal consistency. This is derived by examination of the direction and magnitude of item-scale correlations to see whether items correlate positively ($r \geq 0.4$) with their hypothesised scale. Ware and co-workers[71] suggest that to be satisfactory a minimum of 90% of item-scale correlations should meet the specified criterion; in the present study, item internal consistency was 99%. A second measure of data quality is scaling success, as measured by item discriminant validity. This is assessed by examining the frequency with which items correlate more highly (by 2 standard errors, 95% confidence interval) with their hypothesised scale than with competing scales. At least 80% of item-scale correlations are required to meet this standard;[71] in the present study 83% did so and in all but one the remainder item-scale correlations were higher, though not significantly for their hypothesised, than for competing scales. The final measure of data quality is the percentage of reliable scales, as assessed using Cronbach's coefficient alpha to estimate scale internal consistency. Consistent with previous recommendations, MAP-R for Windows defines as reliable any scale with an alpha coefficient of 0.70 or greater. In the present study, all scales were reliable (100% success) with alpha coefficients ranging from 0.77 (for social functioning) to 0.94 (for health distress).

Descriptive statistics

As for the SF-36, raw scores on MSQOL-54 are linearly transformed into 0–100 scores, where 0 is worse and 100 is better quality of life. *Table 29* compares mean scores and standard deviations for the UK and US samples. Mean scores were similar for both samples, though somewhat lower in the UK sample for bodily pain, energy/vitality and general health perceptions, and higher for satisfaction with sexual function. Examination of floor and ceiling effects (the frequency with which subjects achieve the lowest or highest possible score) is important as marked effects can reduce the ability of an instrument to detect relatively subtle improvement or deterioration, thus limiting its usefulness. As for the US data, there were marked floor effects for two MSQOL-54 scales (role physical, role emotional) with over half of the UK sample scoring the lowest possible score on the scale for role limitations due to physical problems. Similarly, in both US and UK samples marked ceiling effects were observed for two scales (role emotional, bodily pain), with over half of the UK sample scoring the highest possible score on the scale for role limitations due to emotional problems *(Table 29)*.

The above analyses suggest that MSQOL-54 performs similarly in UK and US patients, and its psychometric properties in a UK population are generally acceptable. However, it may be that particular sub-groups complete instruments such as MSQOL-54 less well than others, for reasons of literacy, poorer health, reduced willingness to co-operate and so on. To examine whether the psycho-

TABLE 28 *MSQOL-54 item frequency distributions and numbers missing*

Item	Item frequency distribution										Missing	
	1	2	3	4	5	6	7	8	9	10	Number	%
Physical function												
PF1	84	10	2								6	5.9
PF2	39	42	18								3	2.9
PF3	49	33	15								5	4.9
PF4	69	19	8								6	5.9
PF5	27	50	20								5	4.9
PF6	40	34	23								5	4.9
PF7	80	10	7								5	4.9
PF8	64	17	15								6	5.9
PF9	33	28	35								6	5.9
PF10	10	36	49								7	6.9
Role physical												
RP1	57	38									7	6.9
RP2	74	22									6	5.9
RP3	73	22									7	6.9
RP4	66	30									6	5.9
Role emotional												
RE1	33	60									9	8.8
RE2	41	52									9	8.8
RE3	36	57									9	8.8
Bodily pain												
BP1r	23	17	10	39	8						5	4.9
BP2r	34	26	17	17	4						4	3.9
BP3r*	32	40	15	12	1						2	2.0
General health												
GH1r	1	13	31	42	12						3	2.9
GH2	7	9	27	24	30						5	4.9
GH3r	9	14	15	28	31						5	4.9
GH4	16	30	39	4	8						5	4.9
GH5r	4	23	10	14	47						4	3.9
Energy/vitality												
VT1r	2	14	16	26	15	25					4	3.9
VT2r	0	3	9	28	22	35					5	4.9
VT3	9	25	24	28	11	0					5	4.9
VT4	12	26	31	12	15	1					4	3.9
VT5r*	5	17	10	29	21	15					5	4.9
Social function												
SF1r	18	25	18	27	10						4	3.9
SF2r	24	20	21	25	7						5	4.9
SF3r*	32	26	20	16	6						2	2.0
Mental health												
MH1	3	4	5	20	25	39					6	5.9
MH2	3	6	9	15	24	40					5	4.9
MH3r	6	27	10	36	12	6					5	4.9
MH4	5	7	10	23	29	23					5	4.9
MH5r	14	30	19	21	11	2					5	4.9

r = Original item score reversed so that the higher the score the better the functioning for all items
* *Additional items added to sub-scales of SF-36*

continued

TABLE 28 contd *MSQOL-54 item frequency distributions and numbers missing*

Item	Item frequency distribution										Missing	
	1	2	3	4	5	6	7	8	9	10	Number	%
Cognitive function												
CF1	5	11	11	23	25	23					4	3.9
CF2	4	11	8	20	25	29					4	3.9
CF3	6	11	7	29	28	17					4	3.9
CF4	6	11	7	17	15	39					7	6.9
Health distress												
HD1	7	12	18	22	25	13					5	4.9
HD2	13	19	23	16	22	4					5	4.9
HD3	11	14	12	17	26	17					5	4.9
HD4	9	11	13	18	23	23					5	4.9
Sexual function[†]												
SF1r	20	20	26	16	14						6	5.9
QOL												
QOL1	3	4	2	17	13	16	17	17	4	6	2	2.0
QOL2	3	5	10	44	23	12	2	0	0	0	3	2.9
Change in health												
TRANSr	5	8	45	32	10						2	2.0

r = Original item score reversed so that the higher the score the better the functioning for all items
[†] *Satisfaction with sexual function*

TABLE 29 *Descriptive statistics for MSQOL-54: UK and US samples*

Scale	Number of items	Mean (SD)		% floor		% ceiling	
		USA	UK	USA	UK	USA	UK
Physical function	10	36.7 (32.5)	34.2 (27.0)	13.5	9.2	3.9	2.3
Role physical	4	32.9 (39.0)	28.3 (37.1)	48.6	56.3	17.9	12.6
Role emotional	3	60.0 (42.3)	62.4 (44.5)	26.0	28.7	46.2	54.0
Bodily pain	3	74.2 (25.5)	67.7 (25.4)	1.1	0.0	28.5	23.0
Mental health	5	65.6 (20.4)	67.7 (20.8)	1.1	0.0	1.7	4.6
Energy/vitality	5	42.2 (20.9)	36.6 (18.9)	1.1	3.4	0.6	0.0
Health distress	4	54.4 (26.9)	56.1 (27.9)	2.8	5.7	2.8	2.3
Social function	3	61.7 (25.0)	60.5 (25.6)	1.1	1.1	11.2	10.3
Cognitive function	4	73.0 (24.2)	68.5 (25.6)	1.7	1.1	23.2	9.2
Sexual function[*]	1	50.5 (38.3)	55.5 (31.8)	26.0	11.5	25.3	19.5
General health	5	53.3 (25.3)	41.7 (21.9)	1.2	1.1	2.4	0.0
Overall QOL	2	60.1 (20.1)	55.2 (20.9)	0.6	3.4	2.8	2.3
Change in health	1	46.1 (25.7)	45.1 (23.8)	7.8	10.3	8.4	5.7

[*] *Satisfaction with sexual function*

metric soundness of MSQOL-54 was maintained in specific sub-groups of patients, we re-ran the above analyses for relapse and remission patients separately. For remission patients, data were complete for only 39 (63%) and computable for 52 (84%); for relapse patients, data were complete for 32 (80%) and computable for 36 (90%). The percentages of missing responses across individual items in the scales ranged from 1.3% to 9.1% for remission patients and from 0.0% to 8.3% for relapse patients.

Item internal consistency was 98% for remission patients and 100% for relapse patients. Item discriminant validity was poorer, at 75% and 54%, respectively, reflecting the small sample sizes available for sub-group analyses. However, when the less stringent definition of scaling success was adopted (item-scale correlations greater than item-to-competing-scale correlations, but not necessarily by 2 standard errors) success was 98% in each case. For both sub-groups, all scales had acceptable internal consistency reliability (alpha range: 0.71–0.93 for remission patients, and 0.75–0.95 for relapse patients). Floor effects were particularly problematic for the role limitations due to physical problems scale, where 48% of patients in remission and 77% of patients having had a recent relapse achieved the lowest possible score. Both floor and ceiling effects were also notable for role limitations due to emotional problems in both remission and relapse patients.

Health Technology Assessment 1998; Vol. 2: No. 4 (Update October 1999)

Cost-effectiveness of interferon beta for multiple sclerosis: the implications of new information on clinical effectiveness

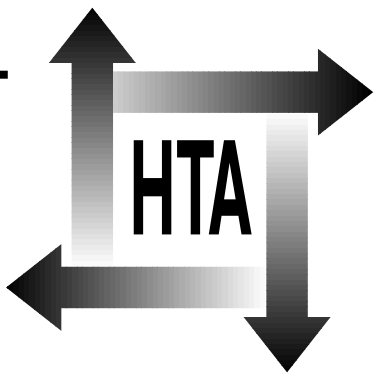
P McNamee*
D Parkin

Department of Epidemiology and Public Health,
School of Health Sciences, University of Newcastle,
Newcastle-upon-Tyne, UK

* Corresponding author

Update *(October 1999)*

Health Technology Assessment
NHS R&D HTA Programme



NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Pharmaceutical Panel and funded as project number 95/01/02 (update 1999).

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein
Monograph Editorial Manager: Melanie Corris

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.

Update:

Cost-effectiveness of interferon beta for multiple sclerosis

Preface

In 1998 the NHS Health Technology Assessment (HTA) programme published a cost–utility analysis of interferon beta for the treatment of relapsing-remitting multiple sclerosis (RRMS) based on evidence from randomised control trials available up to 1997. This monograph gave a detailed account of the methods employed and was, in common with all reports in the HTA series, subject to peer review.

Since publication, important new evidence has become available on the effectiveness of interferon beta in RRMS and, additionally, in secondary progressive disease. The HTA programme has therefore commissioned this short supplement to the 1998 study. As the methods are almost identical, and in the interests of making the results available to the NHS as soon as possible, the supplement has not been peer reviewed.

The NCCHTA welcomes comments on the HTA monograph series; these can be submitted to the HTA programme via its website: http://www.hta.nhsweb.nhs.uk

The Editors
22 July 1999

Background

This paper describes and updates a previous study which was funded by the NHS HTA programme to investigate the cost-effectiveness of interferon beta-1b for the treatment of RRMS.[1]

Within the previous study, a decision analytic model was constructed, based on a series of disease-specific health states (the expanded disability status scale [EDSS]), through which individuals moved over time. Individuals were placed at a particular EDSS level at baseline, and subsequently experienced a certain number of relapses and probability of progression to other EDSS states after a certain number of years. Compared with standard care, therapy was shown to cost £328,300 and £228,300 over 5 and 10 years, respectively, to produce one additional quality-adjusted life year (QALY).

It is possible to explore further the validity of these results, and by implication the potential information provided by further trials, by using other evidence of clinical effectiveness. For example, interferon beta therapies for MS have been recently shown to significantly delay disease progression both for RRMS[2] and secondary progressive disease.[3] However, the magnitude of these effects on health-related quality of life has not been addressed. There are also remaining questions over the magnitude of any cost savings relative to the costs of therapy. To address these issues, data from these trials were used to estimate cost-effectiveness using a similar decision analytic approach.

Methods

The same model was used for disease progression as in the previous NHS HTA study to which readers should refer for full methodological details.[1] In the present study, however, cohorts of individuals were distributed across a series of EDSS levels at baseline according to the trial data, rather than one particular EDSS category, thus improving the model. The proportion of individuals within different cohorts acted as weights in the calculation of transition probabilities. Transition probabilities were calculated by deriving cohort-specific hazard rates for progression and extrapolating over 5- and 10-year periods. Utility values for EDSS states were derived by taking the average EQ-5D tariff for patients in each state. The base year for costs was 1997. The assumptions for the 5- and 10-year models with regard to estimates of probabilities of progression, utilities and costs are summarised in *Tables 1* and *2*.

Results

A series of incremental cost per QALY gained ratios were calculated using 5- and 10-year time periods. Data from the PRISMS (Prevention of Relapses and Disability by Interferon Beta-1a Subcutaneously in Multiple Scelrosis) Study Group, which considered interferon beta-1a for RRMS, produced estimates of £375,100 and £393,300 per

QALY gained over 5 and 10 years, respectively (*Tables 3* and *4*). Higher figures were produced using data from the European Study Group, which considered interferon beta-1b for secondary progressive disease, with cost per QALY gained estimates of £667,800 and £587,200 over 5 and 10 years, respectively (*Tables 5* and *6*).

Conclusions

The values reported above are substantially higher than those calculated in our previous report (£328,300 and £228,300 for 5 and 10 years, respectively). It can be inferred therefore that the validity of our previous conclusions, which stated that therapy produces small short-term gains for large additional costs, are further strengthened as a result of incorporating new trial evidence.

Moreover, due to the nature of reporting in clinical trials of interferon beta-1b, our previous model made a somewhat restrictive assumption by placing all individuals within one particular EDSS category. Recent trials report more disaggregated data; therefore the model now reflects a more realistic situation where individuals are placed within a range of different EDSS levels and experience EDSS-specific transition probabilities.

This type of analysis also shows the extent to which different cohorts are associated with different levels of costs, utilities and QALYs. Although our results suggest that costs are greater and quality of life is lower for individuals with higher EDSS levels, it should not be inferred that treatment will automatically appear less cost-effective for these individuals. Comparison of results from the PRISMS Study Group[2] and the European Study Group[3] shows this to be the case, as exemplified by higher cost–utility ratios in the latter study. However, the determinants are also related to differences in relapse and progression rates between placebo and therapy groups for patients in both studies.

The effect of sensitivity analyses on variation of cost per QALY gained ratios has been demonstrated in the previous report; similar analyses conducted here would have led to similar levels of variation, with no new policy implications. These results, together with previous sensitivity analyses, have implications for further trials as they demonstrate the values required for key variables to produce more acceptable cost–utility ratios. It can be concluded that the results from recent trials do not approach such values by a considerable degree.

References

1. Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D. A cost–utility analysis of interferon beta for multiple sclerosis. *Health Technol Assess* 1998;**2**(4).

2. PRISMS (Prevention of Relapses and Disability by Interferon Beta-1a Subcutaneously in Multiple Sclerosis) Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing-remitting multiple sclerosis. *Lancet* 1998;**352**:1498–504.

3. European Study Group on Interferon Beta-1b 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.

4. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey. University of York: Centre for Health Economics, 1995.

5. Weinshenker BG, Bass B, Rice GPA, Noseworthy J, Carriere W, Baskerville J, *et al.* The natural history of multiple sclerosis: a geographically based study. ii. Predictive value of the early clinical course. *Brain* 1989;**112**:1419–28.

6. Weinshenker BG, Rice GPA, Noseworthy JH, Carriere W, Baskerville J, Ebers GC. The natural history of multiple sclerosis: a geographically based study. iv. Applications to planning and interpretations of clinical therapeutic trials. *Brain* 1991;**114**:1057–67.

TABLE 1 *Assumptions of the PRISMS Study Group model*[2]

Parameter	5-year model		10-year model		Source
	Standard care	IFβ-1a	Standard care	IFβ-1a	
EDSS transition probabilities (weighted by proportion in cohort at baseline)					
1/1.5 → 1/1.5	0.12	0.16	0.06	0.09	PRISMS Study Group[2]
1/1.5 → 2/2.5	0.20	0.16	0.26	0.23	
2/3.5 → 2/3.5	0.20	0.25	0.10	0.14	
2/3.5 → 3/4.5	0.32	0.27	0.42	0.38	
4/5 → 4/5	0.02	0.07	0.00	0.04	
4/5 → 5/6	0.14	0.09	0.16	0.12	
EDSS transition utilities					
1/1.5 → 1/1.5	0.85	0.85	0.85	0.85	Parkin *et al.*[1]; Dolan *et al.*[4]; Weinshenker *et al.*[5,6]
1/1.5 → 2/2.5	0.81	0.81	0.78	0.78	
2/3.5 → 2/3.5	0.76	0.76	0.76	0.76	
2/3.5 → 3/4.5	0.72	0.72	0.69	0.69	
4/5 → 4/5	0.59	0.59	0.59	0.59	
4/5 → 5/6	0.55	0.55	0.52	0.52	
EDSS transition costs per year (£)					
1/1.5 → 1/1.5	375	375	330	330	*Service receipt* – Parkin *et al.*[1] *Unit costs* – CIPFA; NHS Drug Tariff; Newcastle & North Tyne Health Authority; Royal Victoria Infirmary, Newcastle-upon-Tyne
1/1.5 → 2/2.5	515	515	540	540	
2/3.5 → 2/3.5	760	760	665	665	
2/3.5 → 3/4.5	725	725	610	610	
4/5 → 4/5	1080	1080	945	945	
4/5 → 5/6	1205	1205	1135	1135	
Number of relapses	6.4	4.3	9	6	PRISMS Study Group[2]; Weinshenker *et al.*[5,6]
Relapse length (months)	1	1	1	1	Parkin *et al.*[1]
Utility loss per relapse	0.5	0.5	0.5	0.5	Parkin *et al.*[1]
Relapse cost (£)	2115	2115	2115	2115	Parkin *et al.*[1]
IFβ-1a costs per year (£)	0	10,500	0	10,500	NHS Drug Tariff; Parkin *et al.*[1]

CIPFA, Chartered Institute of Public Finance and Accountancy; IFβ, interferon beta

TABLE 2 *Assumptions of the European Study Group model*[3]

Parameter	5-year model		10-year model		Source
	Standard care	IFβ-1b	Standard care	IFβ-1b	
EDSS transition probabilities (weighted by proportion in cohort at baseline)					
3 → 3	0.07	0.08	0.03	0.04	European Study Group[3]
3 → 4	0.09	0.08	0.13	0.12	
4/5 → 4/5	0.13	0.17	0.05	0.08	
4/5 → 5/6	0.26	0.22	0.34	0.31	
6 → 6	0.15	0.21	0.06	0.10	
6 → 7	0.30	0.24	0.39	0.35	
EDSS transition utilities					
3 → 3	0.71	0.71	0.71	0.71	Parkin *et al.*[1]; Dolan *et al.*[4]; Weinshenker *et al.*[5,6]
3 → 4	0.69	0.69	0.67	0.67	
4/5 → 4/5	0.59	0.59	0.59	0.59	
4/5 → 5/6	0.55	0.55	0.52	0.52	
6 → 6	0.49	0.49	0.49	0.49	
6 → 7	0.43	0.43	0.39	0.39	
EDSS transition costs per year (£)					
3 → 3	660	660	570	570	*Service receipt* – Parkin *et al.*[1] *Unit costs* – CIPFA; NHS Drug Tariff; Newcastle & North Tyne Health Authority; Royal Victoria Infirmary, Newcastle-upon-Tyne
3 → 4	700	700	630	630	
4/5 → 4/5	1080	1080	945	945	
4/5 → 5/6	1205	1205	1135	1135	
6 → 6	1420	1420	1240	1240	
6 → 7	1905	1905	1980	1980	
Number of relapses	3.2	2.2	4.5	3	European Study Group[3]; Weinshenker *et al.*[5,6]
Relapse length (months)	1	1	1	1	Parkin *et al.*[1]
Utility loss per relapse	0.5	0.5	0.5	0.5	Parkin *et al.*[1]
Relapse cost (£)	2115	2115	2115	2115	Parkin *et al.*[1]
IFβ-1b costs per year (£)	0	10,500	0	10,500	NHS Drug Tariff; Parkin *et al.*[1]

TABLE 3 *PRISMS Study Group – 5-year decision analytic model*[2]

EDSS	Standard care		IFβ-1a		Costs per QALY gain (£)
	Costs (£)	QALY	Costs (£)	QALY	
1/1.5 → 1/1.5	1925	0.50	9080	0.60	
1/1.5 → 2/2.5	3145	0.74	9565	0.63	
2/3.5 → 2/3.5	3515	0.72	15,245	0.92	
2/3.5 → 3/4.5	5440	1.06	15,820	0.91	
4/5 → 4/5	425	0.06	4425	0.20	
4/5 → 5/6	2690	0.34	5460	0.23	
Total	**17,140**	**3.41**	**59,590**	**3.52**	**375,100**

TABLE 4 *PRISMS Study Group – 10-year decision analytic model*[2]

EDSS	Standard care		IFβ-1a		Costs per QALY gain (£)
	Costs (£)	QALY	Costs (£)	QALY	
1/1.5 → 1/1.5	1430	0.52	8455	0.71	
1/1.5 → 2/2.5	6260	1.90	23,365	1.76	
2/3.5 → 2/3.5	2670	0.75	14,215	1.03	
2/3.5 → 3/4.5	10,450	2.71	38,225	2.52	
4/5 → 4/5	0	0.00	3830	0.21	
4/5 → 5/6	4860	0.77	13,050	0.61	
Total	**25,670**	**6.66**	**101,140**	**6.85**	**393,300**

TABLE 5 *European Study Group – 5-year decision analytic model*[3]

EDSS	Standard care		IFβ-1b		Costs per QALY gain (£)
	Costs (£)	QALY	Costs (£)	QALY	
3 → 3	690	0.23	4535	0.29	
3 → 4	930	0.30	4200	0.26	
4/5 → 4/5	1580	0.37	9845	0.49	
4/5 → 5/6	3370	0.69	12,770	0.59	
6 → 6	2110	0.35	12,085	0.49	
6 → 7	4820	0.60	14,775	0.50	
Total	**13,500**	**2.54**	**58,210**	**2.61**	**667,800**

TABLE 6 *European Study Group – 10-year decision analytic model*[3]

EDSS	Standard care		IFβ-1b		Costs per QALY gain (£)
	Costs (£)	QALY	Costs (£)	QALY	
3 → 3	460	0.21	4185	0.31	
3 → 4	2040	0.85	10,830	0.76	
4/5 → 4/5	970	0.29	8065	0.48	
4/5 → 5/6	7125	1.71	30,920	1.57	
6 → 6	1275	0.27	10,375	0.49	
6 → 7	11,430	1.46	37,280	1.31	
Total	**23,300**	**4.79**	**101,655**	**4.92**	**587,200**

Health Technology Assessment panel membership

This report was identified as a priority by the Pharmaceutical Panel.

Acute Sector Panel

Chair: Professor John Farndon, University of Bristol †

Professor Senga Bond, University of Newcastle-upon-Tyne †

Professor Ian Cameron, Southeast Thames Regional Health Authority

Ms Lynne Clemence, Mid-Kent Health Care Trust †

Professor Francis Creed, University of Manchester †

Professor Cam Donaldson, University of Aberdeen

Mr John Dunning, Papworth Hospital, Cambridge †

Professor Richard Ellis, St James's University Hospital, Leeds

Mr Leonard Fenwick, Freeman Group of Hospitals, Newcastle-upon-Tyne †

Dr David Field, Leicester Royal Infirmary †

Ms Grace Gibbs, West Middlesex University Hospital NHS Trust †

Dr Neville Goodman, Southmead Hospital Services Trust, Bristol †

Mr Ian Hammond, Bedford & Shires Health & Care NHS Trust

Professor Adrian Harris, Churchill Hospital, Oxford

Professor Robert Hawkins, University of Bristol †

Dr Chris McCall, General Practitioner, Dorset †

Professor Alan McGregor, St Thomas's Hospital, London

Mrs Wilma MacPherson, St Thomas's & Guy's Hospitals, London

Professor Jon Nicoll, University of Sheffield †

Professor John Norman, University of Southampton

Dr John Pounsford, Frenchay Hospital, Bristol †

Professor Gordon Stirrat, St Michael's Hospital, Bristol

Professor Michael Sheppard, Queen Elizabeth Hospital, Birmingham †

Dr William Tarnow-Mordi, University of Dundee

Professor Kenneth Taylor, Hammersmith Hospital, London

Diagnostics and Imaging Panel

Chair: Professor Mike Smith, University of Leeds †

Professor Michael Maisey, Guy's & St Thomas's Hospitals, London *

Professor Andrew Adam, UMDS, London †

Dr Pat Cooke, RDRD, Trent Regional Health Authority

Ms Julia Davison, St Bartholomew's Hospital, London †

Professor Adrian Dixon, University of Cambridge †

Professor MA Ferguson-Smith, University of Cambridge †

Dr Mansel Hacney, University of Manchester

Professor Sean Hilton, St George's Hospital Medical School, London

Mr John Hutton, MEDTAP International Inc., London †

Professor Donald Jeffries, St Bartholomew's Hospital, London †

Dr Andrew Moore, Editor, *Bandolier* †

Professor Chris Price, London Hospital Medical School †

Dr Ian Reynolds, Nottingham Health Authority

Professor Colin Roberts, University of Wales College of Medicine †

Miss Annette Sergeant, Chase Farm Hospital, Enfield

Professor John Stuart, University of Birmingham

Dr Ala Szczepura, University of Warwick †

Mr Stephen Thornton, Cambridge & Huntingdon Health Commission

Dr Gillian Vivian, Royal Cornwall Hospitals Trust †

Dr Jo Walsworth-Bell, South Staffordshire Health Authority †

Dr Greg Warner, General Practitioner, Hampshire †

Methodology Panel

Chair: Professor Martin Buxton, Brunel University †

Professor Anthony Culyer, University of York *

Dr Doug Altman, Institute of Health Sciences, Oxford †

Professor Michael Baum, Royal Marsden Hospital

Professor Nick Black, London School of Hygiene & Tropical Medicine †

Professor Ann Bowling, University College London Medical School †

Dr Rory Collins, University of Oxford

Professor George Davey-Smith, University of Bristol

Dr Vikki Entwistle, University of Aberdeen †

Professor Ray Fitzpatrick, University of Oxford †

Professor Stephen Frankel, University of Bristol

Dr Stephen Harrison, University of Leeds

Mr Philip Hewitson, Leeds FHSA

Professor Richard Lilford, Regional Director, R&D, West Midlands †

Mr Nick Mays, King's Fund, London †

Professor Ian Russell, University of York †

Professor David Sackett, Centre for Evidence Based Medicine, Oxford †

Dr Maurice Slevin, St Bartholomew's Hospital, London

Dr David Spiegelhalter, Institute of Public Health, Cambridge †

Professor Charles Warlow, Western General Hospital, Edinburgh †

* Previous Chair
† Current members

continued

continued

Pharmaceutical Panel

Chair: Professor Tom Walley, University of Liverpool †

Professor Michael Rawlins, University of Newcastle-upon-Tyne*

Dr Colin Bradley, University of Birmingham

Professor Alasdair Breckenridge, RDRD, Northwest Regional Health Authority

Ms Christine Clark, Hope Hospital, Salford †

Mrs Julie Dent, Ealing, Hammersmith & Hounslow Health Authority, London

Mr Barrie Dowdeswell, Royal Victoria Infirmary, Newcastle-upon-Tyne

Dr Desmond Fitzgerald, Mere, Bucklow Hill, Cheshire

Dr Alistair Gray, Health Economics Research Unit, University of Oxford †

Professor Keith Gull, University of Manchester

Dr Keith Jones, Medicines Control Agency

Professor Trevor Jones, ABPI, London †

Ms Sally Knight, Lister Hospital, Stevenage †

Dr Andrew Mortimore, Southampton & SW Hants Health Authority †

Mr Nigel Offen, Essex Rivers Healthcare, Colchester †

Dr John Posnett, University of York

Mrs Marianne Rigge, The College of Health, London †

Mr Simon Robbins, Camden & Islington Health Authority, London †

Dr Frances Rotblat, Medicines Control Agency †

Mrs Katrina Simister, Liverpool Health Authority †

Dr Ross Taylor, University of Aberdeen †

Dr Tim van Zwanenberg, Northern Regional Health Authority

Dr Kent Woods, RDRD, Trent RO, Sheffield †

Population Screening Panel

Chair: Professor Sir John Grimley Evans, Radcliffe Infirmary, Oxford †

Dr Sheila Adam, Department of Health*

Ms Stella Burnside, Altnagelvin Hospitals Trust, Londonderry †

Dr Carol Dezateux, Institute of Child Health, London †

Dr Anne Dixon Brown, NHS Executive, Anglia & Oxford †

Professor Dian Donnai, St Mary's Hospital, Manchester †

Dr Tom Fahey, University of Bristol †

Mrs Gillian Fletcher, National Childbirth Trust †

Professor George Freeman, Charing Cross & Westminster Medical School, London

Dr Mike Gill, Brent & Harrow Health Authority †

Dr JA Muir Gray, RDRD, Anglia & Oxford RO †

Dr Ann Ludbrook, University of Aberdeen †

Professor Alexander Markham, St James's University Hospital, Leeds †

Professor Theresa Marteau, UMDS, London

Dr Ann McPherson, General Practitioner, Oxford †

Professor Catherine Peckham, Institute of Child Health, London

Dr Connie Smith, Parkside NHS Trust, London

Dr Sarah Stewart-Brown, University of Oxford †

Ms Polly Toynbee, Journalist †

Professor Nick Wald, University of London †

Professor Ciaran Woodman, Centre for Cancer Epidemiology, Manchester

Primary and Community Care Panel

Chair: Dr John Tripp, Royal Devon & Exeter Healthcare NHS Trust †

Professor Angela Coulter, King's Fund, London *

Professor Martin Roland, University of Manchester *

Dr Simon Allison, University of Nottingham

Mr Kevin Barton, East London & City Health Authority †

Professor John Bond, University of Newcastle-upon-Tyne †

Ms Judith Brodie, Age Concern, London †

Dr Nicky Cullum, University of York †

Professor Shah Ebrahim, Royal Free Hospital, London

Mr Andrew Farmer, Institute of Health Sciences, Oxford †

Ms Cathy Gritzner, The Patients' Association †

Professor Andrew Haines, RDRD, North Thames Regional Health Authority

Dr Nicholas Hicks, Oxfordshire Health Authority †

Professor Richard Hobbs, University of Birmingham †

Professor Allen Hutchinson, University of Sheffield †

Mr Edward Jones, Rochdale FHSA

Professor Roger Jones, UMDS, London

Mr Lionel Joyce, Chief Executive, Newcastle City Health NHS Trust

Professor Martin Knapp, London School of Economics & Political Science

Professor Karen Luker, University of Liverpool

Professor David Mant, NHS Executive South & West †

Dr Fiona Moss, North Thames British Postgraduate Medical Federation †

Professor Dianne Newham, King's College London

Professor Gillian Parker, University of Leicester †

Dr Robert Peveler, University of Southampton †

Dr Mary Renfrew, University of Oxford

Ms Hilary Scott, Tower Hamlets Healthcare NHS Trust, London †

* Previous Chair
† Current members

Health Technology Assessment 1998; Vol. 2: No. 4

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.hta.nhsweb.nhs.uk

ISSN 1366-5278