



## RESEARCH PAPER

# Assessing the long-term effectiveness of interferon-beta and glatiramer acetate in multiple sclerosis: final 10-year results from the UK multiple sclerosis risk-sharing scheme

Jacqueline Palace,<sup>1</sup> Martin Duddy,<sup>2</sup> Michael Lawton,<sup>3</sup> Thomas Bregenzer,<sup>4</sup> Feng Zhu,<sup>5</sup> Mike Boggild,<sup>6</sup> Benjamin Piske,<sup>4</sup> Neil P Robertson,<sup>7</sup> Joel Oger,<sup>5</sup> Helen Tremlett,<sup>5</sup> Kate Tilling,<sup>3</sup> Yoav Ben-Shlomo,<sup>3</sup> Richard Lilford,<sup>8</sup> Charles Dobson<sup>9</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jnnp-2018-318360>).

For numbered affiliations see end of article.

## Correspondence to

Dr Jacqueline Palace, Clinical Neurology, The Oxford University Hospitals Trust, Oxford OX3 9DU, UK; [jacqueline.palace@ndcn.ox.ac.uk](mailto:jacqueline.palace@ndcn.ox.ac.uk)

YB-S and RL contributed equally.

Received 6 March 2018  
Revised 14 June 2018  
Accepted 7 July 2018

## ABSTRACT

**Background** Because multiple sclerosis (MS) is a chronic disease causing disability over decades, it is crucial to know if the short-term effects of disease-modifying therapies reported in randomised controlled trials reduce long-term disability. This 10-year prospective observational study of disability outcomes (Expanded Disability Status Scale (EDSS) and utility) was set up, in conjunction with a risk-sharing agreement between payers and producers, to investigate this issue.

**Methods** The outcomes of the UK treated patients were compared with a modelled untreated control based on the British Columbia MS data set to assess the long-term effectiveness of these treatments. Two complementary analysis models were used: a multilevel model (MLM) and a continuous Markov model.

**Results** 4862 patients with MS were eligible for the primary analysis (mean and median follow-up times 8.7 and 10 years). EDSS worsening was reduced by 28% (MLM), 7% (Markov) and 24% time-adjusted Markov in the total cohort, and by 31% (MLM) and 14% (Markov) for relapsing remitting patients. The utility worsening was reduced by 23%–24% in the total cohort and by 24%–31% in the RR patients depending on the model used. All sensitivity analyses showed a treatment effect. There was a 4-year (CI 2.7 to 5.3) delay to EDSS 6.0. An apparent waning of treatment effect with time was seen. Subgroup analyses suggested better treatment effects in those treated earlier and with lower EDSS scores.

**Conclusions** This study supports a beneficial effect on long-term disability with first-line MS disease-modifying treatments, which is clinically meaningful. However the waning effect noted requires further study.

reported in the 1990s.<sup>2–5</sup> These led to the licensing of three formulations of interferon beta (Betaferon/Betaseron, Rebif and Avonex) and of glatiramer acetate (Copaxone) for RRMS. The licences of Betaferon and Rebif were subsequently extended in Europe to include patients with relapsing SPMS.<sup>6–8</sup>

All four treatments were shown, over the period of 2–3 years of the trials, to be efficacious in reducing relapses (by approximately 30%) and in reducing MRI activity, with less robust evidence on disability. The most important outcome, that is, the long-term effect of treatment on disability, could not be addressed within this timescale, and the predictive value of short-term treatment outcomes on longer term disability remains unproven. Only longer term follow-up data can provide information on the sustainability of treatment effects and on delaying time to loss of independent ambulation.

Because of the uncertainty of the long-term benefit, the UK's National Institute for Health and Care Excellence (NICE) in 2002 concluded that it was unable to recommend the use of these DMTs within the UK National Health Service,<sup>9</sup> but recognised that these drugs could be cost-effective over the longer term if early treatment effects on disability persisted. Consequently a novel risk-sharing scheme (RSS) was set up between the UK Department of Health, the pharmaceutical companies, professional and patient groups to deliver these DMTs cost-effectively<sup>10</sup> (see online supplementary appendix 1 for details).

A key feature of the scheme was the monitoring of disability progression in a cohort of patients to test whether observed outcomes were in line with those required for cost-effectiveness. This observational study recruited over 5000 UK patients prescribed the interferon-βs and glatiramer acetate between 2002 and 2005 and followed them up over a 10-year period in order to measure the long-term effectiveness of the drugs when compared with a modelled natural history cohort. Two yearly analyses were performed and if observed outcomes deviated from target by more than an agreed margin, the price of the drugs would be adjusted to restore cost-effectiveness.

At 2 years the initial Ontario natural history data set was deemed not suitable to model the untreated

## INTRODUCTION

Multiple sclerosis (MS) is a major cause of serious physical disability in adults of working age. The majority of patients start with a relapsing remitting phase (RRMS), which then becomes secondary progressive (SPMS) with or without superimposed relapses at a median disease duration of 20 years.<sup>1</sup> It is during this latter phase that the majority of the disability is manifest.

The first randomised controlled trials (RCT) of MS disease-modifying treatments (DMTs) were



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Palace J, Duddy M, Lawton M, et al. *J Neurol Neurosurg Psychiatry* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jnnp-2018-318360

## Multiple sclerosis

control group because the disability scores had been smoothed,<sup>11</sup> and the British Columbia MS database was selected and validated.<sup>12–13</sup> A second independent analysis model (a multilevel model (MLM)) was added<sup>14</sup> to corroborate results from the Markov model specified in the RSS.<sup>15</sup>

After 6 years of follow-up,<sup>13</sup> the drugs in aggregate showed a 40% reduction in the rate of deterioration in disability on the MLM in patients with relapsing remitting disease (24% on the Markov model); this translated into a reduction of 43% (multilevel) or 42% (Markov) in the rate at which utility worsened, which was on track for the cost-effective target.<sup>10</sup>

The scientific analysis plan for this 10-year analysis<sup>16</sup> was further revised in line with intention-to-treat principles, with additional analyses of (1) subgroups, (2) the influence of time on treatment effectiveness and (3) the effect of treatment on the time to loss of unaided ambulation. Here we report the final results of this study, focusing on the longer term effect of DMTs on disability progression.

### METHODS

The detailed protocol can be found elsewhere<sup>12–14</sup> and the following is a summary.

#### Patient recruitment and follow up

Between May 2002 and July 2005, 72 sites across the UK recruited patients with RRMS or SPMS fulfilling the Association of British Neurologists (ABN) 2001 criteria for treatment (Expanded Disability Status Scale (EDSS)  $\leq 6.5$ ;  $\geq 18$  years old; two relapses in the last two calendar years).<sup>17</sup> Drug selection reflected clinical practice and was led by individual patient and physician choice within the licensed indications in the UK. Ten-year follow-up was planned with the annual EDSS<sup>18</sup> scores, whether treatment was continued or not. Telephone assessments were permitted for EDSS values over 6.<sup>19</sup>

For this final analysis the main outcome was an ‘intention to treat’ analysis which included all recruited patients, both RRMS and SPMS at baseline, and all follow-up scores including those from patients switching to non-scheme DMTs. However, subgroup analysis also focused on the RRMS at baseline patients, since this was the cohort for which the price adjustment scheme was set up.

#### Natural history comparator

The British Columbia MS data set is a population-based database, established in 1980, with the EDSS scored by the MS neurologist at outpatient visits.<sup>1</sup> Patients (n=978) were identified who, between 1980 and 1995, fulfilled similar eligibility criteria to the RSS cohort. The ‘baseline’ for these patients was taken as the first clinic visit at which they fulfilled the criteria. Patient information was included only up to the end of 1995, after which DMTs were widely available in Canada.

#### Outcomes

For scientific purposes, our primary interest was to compare the rate of disability worsening (ie, disease progression) of patients treated with DMTs, as observed in the RSS, with that in an untreated modelled comparator control group (‘comparison against control’). Disease progression was modelled in the patients in the British Columbia data set adjusting for differences in prognostic factors (EDSS at baseline, age at onset) between the two data sets. The technical details are explained below. Disease progression was measured in terms of the accumulation of disability (as measured by EDSS) and worsening of quality of life

(expressed as utility, derived from the EDSS score—see online supplementary appendix 2), for eligible and treated patients with at least one EDSS score after baseline. Outcomes were expressed both as absolute differences between progression observed in the RSS treated group and the comparator control group, and the relative difference in progression (expressed as a percentage).

The secondary outcomes included the area under the curve (a cumulative measure of reduction in disability using the absolute value of area under EDSS–time curve calculated using the trapezoid rule) and the delay in the median time required for sustained progression to the clinically relevant milestone of EDSS 6.0 (needing a stick to walk 100 m), confirmed at least 6 months with no subsequent lower scores.<sup>20</sup>

#### Supplementary analyses

A wide range of prespecified sensitivity analyses (online supplementary table 1 and 2) were conducted to assess assumptions related to modelling and loss to follow-up. Prespecified subgroup analyses allowed us to examine treatment effect by sex, EDSS and disease course (RRMS/SPMS) at baseline, disease duration and recruitment date. (Patients recruited later in the scheme may be more typical of normal clinical practice than those recruited earlier, many of whom had been waiting for treatment and had more advanced disease.) We also used a variety of methods to explore the apparent changes in the treatment effect over time (online supplementary appendix 3).

#### Statistical design

In order to calculate the expected progression for untreated patients, two independent models were developed, a Markov model and an MLM—for further details see online supplementary appendix 4. Both models used data from the natural history comparator data set and these were validated in previous studies<sup>12–14</sup> as being reliable methods for predicting the untreated outcome. The essential role of the modelling is to adjust for the (relatively slight) differences in baseline characteristics between the natural history and RSS data sets.

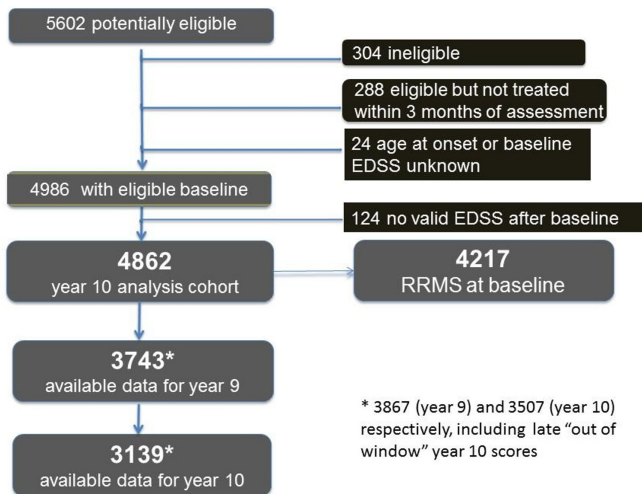
#### Markov model

The Markov model uses the annual probabilities of moving between EDSS scores (or remaining on the same score) from the untreated values derived from the British Columbia data set. These probabilities are then applied to the baseline EDSS scores of the RSS cohort and to subsequent modelled EDSS states over 10 years to give the expected EDSS progression for untreated patients. The difference between the expected untreated and observed RSS treated mean EDSS score represents the treatment effect.

Estimation was by the Jackson continuous time method<sup>21</sup> using age at onset dichotomised at the median as a predetermined covariate.<sup>12</sup> Because the assumption in the Markov model is that the transition probabilities from one specific EDSS score do not vary over time (and this assumption is unlikely to be precise), we used a ‘time-varying’ model, with separate transition matrices estimated for the first 2 years after baseline and for the rest of the follow-up period, as a sensitivity analysis. CIs on projections using the Markov model are derived by bootstrapping.

#### Multilevel model

The MLM<sup>22–23</sup> uses the mean trajectory for the whole population, the variation of individual trajectories about this mean and the fluctuation of individual EDSS scores about the trend for each individual. The untreated values were again derived



**Figure 1** Consolidated Standards of Reporting Trials diagram for all patients in the UK MS risk-sharing scheme. EDSS, Expanded Disability Status Scale; MS, multiple sclerosis; RRMS, relapsing remitting MS.

from the British Columbia data set. The model is applied to the baseline data for patients in the RSS data set to predict the EDSS progression without treatment for each individual. The projections are then combined across all individuals to produce a predicted untreated mean EDSS progression for the whole RSS cohort. This is then compared with the observed data in the RSS (the 'comparison against control') to estimate the absolute treatment effect.

For consistency with the Markov model,<sup>12</sup> age at onset (binary variable) was included as the only covariate after previous covariate modelling found this to be sufficient. Validation of the model<sup>14 24</sup> was undertaken on the British Columbia cohort (internal validation) and also using a natural history data set from the University Hospital of Wales (external validation). Bootstrapping was used to derive 95% CIs.

### Time to sustained EDSS 6.0

To estimate the time to sustained EDSS 6.0, a parametric (Weibull) model was fitted to the British Columbia Multiple Sclerosis (BCMS) data, with gender and baseline EDSS as covariates. A similar model was fitted to the RSS data. To adjust for differences in baseline distributions, the strata were combined using as weights the proportion of patients at baseline in the RSS data set (see online supplementary appendix 5 for details).

### Governance of the study

An independent Scientific Advisory Group consisting of Pelham Barton, Yoav Ben-Shlomo, Richard Gray and chaired by Richard Lilford developed and approved the statistical analysis plan. Representation from the authors, NICE, the MS Trust and the Department of Health attended the Scientific Advisory Group meetings. The representatives from each company were observers at the Scientific Advisory Group meetings but had no role in the data collection, analysis or preparation of the manuscript.

## RESULTS

### Patient disposition and characteristics

Out of 5602 patients registered for recruitment, 4986 were consented and eligible for treatment, began treatment, and had baseline assessments. Of these 4986, 4862 (97.5%) had at least one follow-up EDSS score and were included in the primary

analysis. Of these, 17% switched to non-scheme drugs during follow-up, with the majority using natalizumab or fingolimod, and 18% discontinued treatment during the study. The Consolidated Standards of Reporting Trials diagram is shown in figure 1 (details in online supplementary table 2). The median follow-up was 10 years, the mean follow-up was 8.7 years and 77% of the primary analysis group had a minimum of 9 years of follow-up data. An additional 2.6% had late year 10 data collected, used in the sensitivity analysis.

The baseline characteristics for the analysis cohort were similar to those in the British Columbia data set (online supplementary table 3), although RSS patients had on average slightly higher EDSS and age at onset and thus slightly longer disease duration on first assessment; however, the models adjust for baseline EDSS scores and age of onset.

### Primary outcomes

The primary outcomes are shown in table 1. The absolute 10-year treatment effect was a reduction of 0.61 using the MLM (95% CIs 0.55 to 0.66) and of 0.12 in mean EDSS using the Markov model (95% CIs 0.07 to 0.17). The corresponding relative reduction in progression was 28% (95% CIs 26% to 31%) for the MLM and 7% (95% CIs 4% to 10%) for the Markov model. However, using the time-varying Markov model, the effect size closely mirrors the MLM results (see below). In utility terms, the reduction in progression was 23% (95% CIs 20% to 26%) on the MLM and 24% (95% CIs 21% to 27%) on the Markov model. Figure 2A,B shows how the observed disability progression deviates over the 10-year period from that expected for untreated patients (the comparator control group) for the MLM.

### Secondary outcomes

The cumulative benefit accrued over the first 10 years of treatment for the whole cohort (the area under the curve for the full period) was 1.3 EDSS years (95% CIs 0.9 to 1.6—estimates available only for the Markov model). The estimated median time for the whole cohort to reach sustained EDSS 6.0 using the Weibull model was 12.5 years (95% CIs 11.8 to 13.3); this compares with 8.4 years (95% CIs 7.4 to 9.6) for untreated patients (figure 3), indicating that treatment is associated with a delay of 4.0 years (95% CIs 2.7 to 5.3) in reaching this relevant disability endpoint.

### Sensitivity analyses

The results of the sensitivity analyses are shown in table 2 and online supplementary table 4. As expected, excluding EDSS scores after patients switched to alternative treatments—effectively censoring later data from patients on a worse trajectory—makes the results appear more favourable to the DMTs. Imputing missing values in the RSS data set makes relatively little difference to the outcomes, except in the deliberately extreme case (greater than a standard worse case scenario) in which we added 0.5 EDSS points to each imputed value (on average, doubling the imputed progression since the last available score), where the magnitude of the treatment effect on EDSS was reduced by roughly a quarter. In contrast, a variant in which we 'enriched' the British Columbia data set by imputing additional values for patients with relatively sparse follow-up and re-estimated the transition probabilities resulted in a substantially greater treatment effect. Finally, using variant of the Markov model with time-varying transition probabilities resulted in a greater treatment effect, with a much closer agreement for the EDSS results

**Table 1** Outcomes of the primary analysis: primary analysis cohort, including patients with SPMS at baseline (n=4862; average follow-up 8.7 years)

Outcome measure	Model	Actual progression (95% CI)	Predicted progression (natural history) (95% CI)	Absolute treatment effect (predicted natural history progression less actual) (95% CI)	Relative rate of disease progression (actual divided by predicted natural history) (95% CI)
		(1)	(2)	(3)=(2-1)	(4)=(1)/(2)
Rounded EDSS*	Markov	1.53 (1.48 to 1.58)	1.65 (1.63 to 1.67)	0.12 (0.07 to 0.17)	93% (90% to 96%)
EDSS*	MLM	1.53 (1.47 to 1.58)	2.13 (2.11 to 2.15)	0.61 (0.55 to 0.66)	72% (69% to 74%)
Utility	Markov	0.122 (0.117 to 0.127)	0.161 (0.159 to 0.163)	0.039 (0.034 to 0.044)	76% (73% to 79%)
	MLM	0.122 (0.117 to 0.127)	0.159 (0.156 to 0.162)	0.037 (0.031 to 0.042)	77% (74% to 80%)

\*For the Markov model, half-integer EDSS states are combined with the next lower integer EDSS state. Calculations with the MLM using rounded EDSS values gave almost identical results to those using the full EDSS scale.  
EDSS, Expanded Disability Status Scale; MLM, multilevel model; SPMS, secondary progressive multiple sclerosis.

to those from the MLM. Online supplementary appendices 6 and 7 provide further details.

**Subgroup analyses**

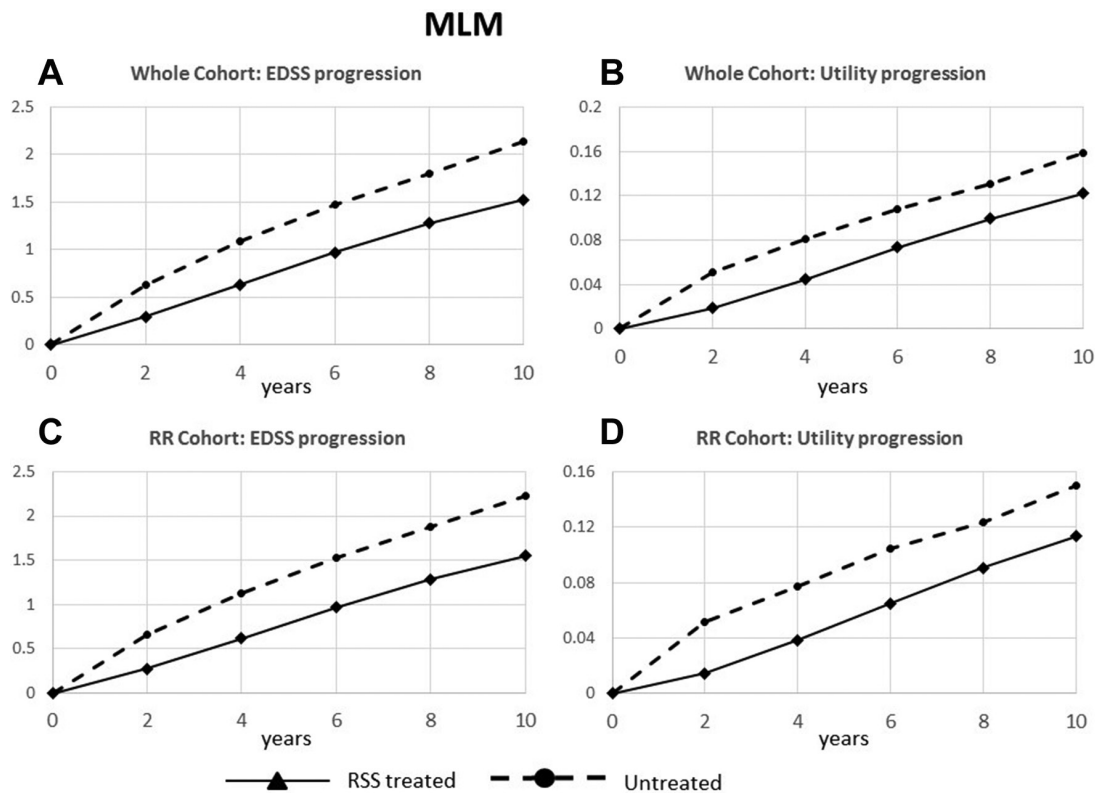
Estimated treatment effects were rather larger for the subset of patients with RRMS at baseline (table 3, figure 2C,D; MLM), with 31% and 14% reduction in EDSS worsening and 24% and 31% reduction in utility worsening (MLM and Markov models, respectively).

Other prespecified subgroup analyses using the Markov model (tables 4 and 5, and online supplementary table 5) suggest that greater treatment effects are associated with a lower EDSS score at baseline, patients with shorter time from disease onset, women rather than men, patients on treatment throughout rather than those who come off treatment, and patients recruited later in the

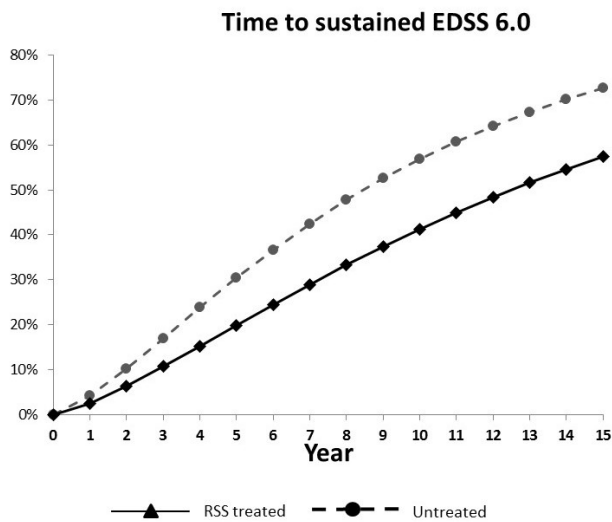
scheme. The MLM however did not find evidence of a reduction in treatment effect with increasing baseline EDSS.

**Changes over time**

Table 6 shows how the main outcomes using the MLM model change when the primary analysis is repeated with years 2, 4, 6, 8 and 10 as the end year. The results suggest that, although the treatment effect continues throughout the follow-up period, it becomes progressively smaller. We explored this through a number of prespecified supplementary analyses using both models, including the version of the Markov model with time-variant transition probabilities. The results (online supplementary appendix 8) confirm this trend of a large initial effect—in the first year or two after treatment initiation—followed by a smaller continuing effect. A sensitivity analysis using year 1 as



**Figure 2** Comparison against control results using the MLM model: disability outcomes over 10 years in RSS patients and in the untreated comparator control group, for the whole cohort (A) EDSS progression and (B) utility progression, and for the patients with RRMS at baseline (C) EDSS progression and (D) utility progression. EDSS, Expanded Disability Status Scale; MLM, multilevel model; RRMS, relapsing remitting multiple sclerosis; RSS, risk-sharing scheme.



**Figure 3** Time to sustained EDSS 6.0 in RSS treated patients and in the untreated comparator control group. EDSS, Expanded Disability Status Scale; RSS, risk-sharing scheme.

the baseline (table 2 and online supplementary table 4) also showed a positive but smaller treatment effect.

## DISCUSSION

### Main findings

Our findings show a clinically significant treatment effect maintained at 10 years, on both EDSS and quality of life outcomes, with a meaningful delay to sustained loss of unaided ambulation of around 4 years in an already impaired cohort at treatment onset (median EDSS 3.5).

Our analysis also highlighted a gradual attenuation of the treatment effect over time, or perhaps a large initial effect followed by a more modest continuing effect. The treatment effects appeared larger for patients with RRMS, those who start treatment earlier in their disease course and for women. The evidence for the effect of baseline EDSS on the treatment effect is less clear.

### Methodological issues

This unique study addresses an important issue related to the long-term impact of the first-generation DMTs, which was not possible to address by a conventional RCT. By establishing a

**Table 2** Sensitivity analyses (EDSS basis)

Variant	Patients (n)	Mean years		Absolute treatment effect (EDSS)		Relative treatment effect (EDSS)	
		Follow-up	On treatment	Markov model	MLM	Markov model	MLM
Primary analysis	4862	8.7	7.0	0.12 (0.07, 0.17)	0.61 (0.55, 0.66)	93% (90%, 96%)	72% (69%, 74%)
RRMS-only subgroup	4217	8.9	7.3	0.25 (0.20, 0.31)	0.68 (0.62, 0.74)	86% (83%, 89%)	69% (67%, 72%)
Supplementary/sensitivity analyses (including patients with SPMS at baseline except where noted)							
Excluding EDSS scores recorded after a patient has switched to a non-scheme DMT*							
Primary analysis cohort	4799	8.0	6.4	0.13 (0.08, 0.19)	0.59 (0.53, 0.64)	91% (88%, 95%)	70% (68%, 73%)
RRMS-only subgroup†	4157	8.1	6.7	0.26 (0.21, 0.32)	0.66 (0.60, 0.72)	84% (81%, 88%)	68% (65%, 71%)
Excluding EDSS scores recorded after a patient has switched to any other DMT*							
Primary analysis cohort	4475	7.0	5.7	0.16 (0.11, 0.22)	0.57 (0.52, 0.62)	88% (84%, 92%)	68% (65%, 70%)
RRMS-only subgroup	3871	7.0	5.9	0.29 (0.23, 0.35)	0.64 (0.59, 0.70)	80% (76%, 84%)	65% (62%, 68%)
Imputation—Markov model, using imputed values derived from the MLM							
'on treatment' assumption	4862	9.9	7.7	0.12 (0.07, 0.17)	NA	93% (91%, 96%)	NA
'off treatment' assumption	4862	9.9	7.0	0.10 (0.04, 0.15)	NA	95% (92%, 98%)	NA
Imputation—multilevel model							
'on treatment' assumption	4862	9.9	9.9	NA	0.67 (0.62, 0.72)	NA	72% (70%, 74%)
'off treatment' assumption	4862	9.9	9.9	NA	0.65 (0.59, 0.70)	NA	73% (70%, 75%)
Multiple imputation	4862	9.9	9.9	NA	0.64 (0.59, 0.70)	NA	73% (71%, 75%)
Single imputation +0.5 EDSS points for each imputed value	4862	9.9	9.9	NA	0.48 (0.43, 0.53)	NA	80% (78%, 82%)
Supplement RSS data with imputed values derived from out-of-window year 10 scores‡	4862	8.9	7.0	0.12 (0.07, 0.18)	‡	93% (90%, 96%)	‡
Use transition matrices from BCMS data set supplemented by imputing additional data for patients with sparse follow-up‡							
Including patients with SPMS at baseline	4862	8.7	7.0	0.28 (0.23, 0.34)	‡	84% (81%, 87%)	‡
Excluding patients with SPMS at baseline	4217	8.9	7.3	0.43 (0.37, 0.49)	‡	78% (75%, 81%)	‡
Alternative Markov model with time-varying natural history transition matrices	4862	8.7	7.0	0.48 (0.42, 0.53)	NA	76% (74%, 79%)	NA
Year 1 baseline (see online supplementary appendix 8)	4360	7.9	6.2	0.07 (0.02, 0.12)	0.43 (0.38, 0.48)	95% (92%, 99%)	77% (74%, 79%)

\*Patients switching before year 1 are also excluded from the analysis.

†This was the primary analysis for the interim year 6 analysis.

‡These variants were only carried out for the Markov model.

DMT, disease-modifying treatment; EDSS, Expanded Disability Status Scale; MLM, multilevel model; NA, not applicable; RRMS, relapsing remitting multiple sclerosis; RSS, risk sharing scheme; SPMS, secondary progressive multiple sclerosis; BCMS, British Columbia Multiple Sclerosis.

**Table 3** Main outcomes, patients with RRMS at baseline: subgroup of patients with RRMS at baseline (n=4217; average follow-up 8.9 years)

Outcome measure	Model	Actual progression (95% CI)	Predicted progression (natural history)(95% CI)	Absolute treatment effect (predicted natural history progression less actual) (95% CI)	Relative rate of disease progression (actual divided by predicted natural history)(95% CI)
		(1)	(2)	(3)=(2-1)	(4)=(1)/(2)
Rounded EDSS*	Markov	1.55 (1.49 to 1.61)	1.80 (1.78 to 1.83)	0.25 (0.19 to 0.31)	86% (83% to 89%)
EDSS*	MLM	1.55 (1.49 to 1.61)	2.23 (2.21 to 2.25)	0.68 (0.62 to 0.74)	69% (67% to 72%)
Utility	Markov	0.113 (0.108 to 0.119)	0.164 (0.163 to 0.166)	0.051 (0.046 to 0.056)	69% (66% to 72%)
	MLM	0.113 (0.108 to 0.118)	0.150 (0.148 to 0.152)	0.037 (0.031 to 0.042)	76% (72% to 79%)

\*For the Markov model, half-integer EDSS states are combined with the next lower integer EDSS state. Calculations with the MLM using rounded EDSS values gave almost identical results to those using the full EDSS scale.

EDSS, Expanded Disability Status Scale; MLM, multilevel model; RRMS, relapsing remitting multiple sclerosis.

novel RSS between the manufacturers and payers, we were able to collect high-quality long-term data on disease progression on a treated cohort and use historical data on untreated patients to construct a comparison control group.

Given the inevitable limitations of observational data, our study has important strengths. The RSS cohort is one of the largest cohorts ever used in a study of this type, with over 42 000 patient years of data, with 9 years or more of data for nearly 80% of patients—a dropout rate over a decade comparable with that seen in recent 2-year treatment trials.<sup>25 26</sup> Moreover these results reflect the ‘real world’ performance of the DMTs and include the effects of patients discontinuing treatment, and may be more generalisable than estimates from highly selected short-term RCT populations. Our natural history comparator, from the British Columbia data set, is also the largest and most complete data set of untreated patients available for MS research.<sup>1</sup> We used two completely independent modelling techniques, with complementary strengths and weaknesses: the multilevel repeated measures model can more easily take account of variations in the rate of disease progression between individual disease trajectories, whereas the Markov model predicted more accurately the number of patients at high EDSS levels which largely determine changes in mean utility (see (online supplementary appendix 4). A third model, used for the survival analysis, confirmed the finding of a clinically significant treatment effect on time to loss of unaided ambulation.

We have implemented ‘intention to treat’ principles, adapting the methods of our 6-year study<sup>13</sup> to include all eligible and consented patients with MS prescribed these DMTs, including those with SPMS as well as RRMS at baseline, and data from patients who switched to different DMTs—including non-scheme DMTs—during follow-up. We believe therefore that we have mitigated an important source of potential bias. Naturally, the possibility of bias from incomplete follow-up remains, but we have tested this by extensive sensitivity analyses with imputation of missing values. In addition, the patterns of follow-up seem broadly similar between the RSS and the British Columbia cohorts,<sup>13</sup> so any residual biases are likely to offset one another (see online supplementary appendix 7). If anything, our results may be biased against a treatment effect because we used telephone questionnaires in the UK cohort to capture data from patients with high EDSS scores who were unable to attend clinic, whereas there was no equivalent for the BCMS cohort. Our models are designed to adjust for the (relatively small) difference in baseline characteristics between the British Columbia and RSS data sets, but if this adjustment is imperfect this would be expected to bias the results *against* the treatment effect since RSS patients had on average marginally worse prognostic factors than British Columbia patients.

There are several other potential biases that must be considered before assuming a causal explanation for the DMTs that may have resulted in differences in the natural history of the two cohorts unrelated to treatment.

### Geographical differences

Could our UK patients have a milder form of MS than the natural history British Columbia patients? Although British Columbia is more ethnically diverse, MS still predominates in Europeans there.<sup>27</sup> More direct evidence demonstrated similar untreated trajectories by using our Canadian data set to reproduce accurately 10-year outcomes in a Welsh untreated MS cohort.<sup>24</sup>

### Temporal differences

Could untreated patients with MS have a slower disease progression than they used to? This concept has been supported by the observation that the relapse rates in the placebo arms of RCTs have been reducing. This is likely to be due to other reasons however. Now DMTs are widely available, ethical considerations make it likely that patients with more active disease are prescribed MS drugs rather than risk being allocated to a placebo arm. Additionally the later trials introduced early rescue therapy options and stricter relapse definitions, which would also reduce relapse rates recorded.<sup>28</sup> Population-based cohorts are less likely to be biased by selection, and there is no evidence for a better prognosis in recent MS cohorts from British Columbia<sup>29</sup> nor Wales (Neil Robertson, personal communication).

### Confounding by indication

Could treatment decisions and clinical outcomes be confounded by variables such as disease severity and comorbidity, as seen in contemporary treated and untreated comparisons? This is not relevant in this study because there was no treated versus untreated comparison from within a single cohort. The BCMS patients were selected using the same inclusion criteria applied to the UK patients. Additionally both analysis methods adjusted for baseline EDSS and age at entry, and adding other covariates made little difference to the outcomes.<sup>12</sup> Previous results from the British Columbia data<sup>20 30</sup> demonstrated that indication bias may be important when comparing with contemporary untreated patients,<sup>20</sup> because untreated patients differed, having a milder disease even though fulfilling the same baseline criteria for treatment. When the authors compared treated patients with historical controls as in our study, they showed a trend towards a treatment benefit on time to reach EDSS 6.0 (relative risk 0.77, CI 0.58 to 1.02), which is comparable with our results considering they had a much smaller treated cohort (n=868 vs 4986) with shorter follow-up time (mean 5.2 and median 5.1 years vs

**Table 4** Subgroup analyses by type of MS, date of baseline assessment and baseline EDSS (both models): EDSS outcomes

Population	Mean years			Mean		Markov model		Multilevel model		Multilevel model with imputation	
	n	At risk	On treatment	Baseline EDSS	Absolute treatment effect	Relative rate of progression	Absolute treatment effect	Relative rate of progression	Absolute treatment effect	Relative rate of progression	Relative rate of progression
Primary analysis population	4862	8.7	7.0	3.18	0.12 (0.07, 0.17)	93% (90%, 96%)	0.61 (0.55, 0.66)	72% (69%, 74%)	0.65 (0.59, 0.70)	73% (70%, 75%)	73% (70%, 75%)
Baseline EDSS ≤3.5	2940	9.0	7.4	2.02	0.35 (0.28, 0.42)	84% (81%, 87%)	0.63 (0.56, 0.71)	74% (71%, 77%)	0.67 (0.59, 0.74)	75% (73%, 78%)	75% (73%, 78%)
Baseline EDSS 4–5.5	1275	8.6	6.7	4.42	-0.11 (-0.21, -0.02)	110% (102%, 119%)	0.59 (0.49, 0.68)	68% (63%, 73%)	0.64 (0.54, 0.73)	69% (65%, 73%)	69% (65%, 73%)
Baseline EDSS ≥6	647	8.0	5.5	6	-0.47 (-0.57, -0.36)	233% (202%, 265%)	0.52 (0.41, 0.62)	63% (56%, 70%)	0.59 (0.48, 0.69)	65% (59%, 71%)	65% (59%, 71%)
RRMS at baseline	4217	8.9	7.3	2.86	0.25 (0.20, 0.31)	86% (83%, 89%)	0.68 (0.62, 0.74)	69% (67%, 72%)	0.71 (0.66, 0.78)	71% (69%, 73%)	71% (69%, 73%)
SPMS at baseline	645	7.8	5.1	5.29	-0.76 (-0.86, -0.66)	218% (202%, 234%)	0.10 (0.00, 0.20)	93% (86%, 101%)	0.19 (0.09, 0.29)	90% (85%, 95%)	90% (85%, 95%)
RRMS at baseline and:											
Baseline EDSS ≤3.5	2882	9.0	7.5	2.00	0.39 (0.32, 0.46)	82% (79%, 86%)	0.67 (0.59, 0.74)	73% (69%, 76%)	0.70 (0.63, 0.77)	74% (71%, 77%)	74% (71%, 77%)
Baseline EDSS 4–5.5	1093	8.7	7.0	4.40	-0.01 (-0.11, 0.10)	101% (91%, 110%)	0.68 (0.58, 0.79)	63% (58%, 68%)	0.72 (0.62, 0.83)	65% (60%, 70%)	65% (60%, 70%)
Baseline EDSS ≥6	242	8.4	6.2	6	-0.16 (-0.35, 0.04)	143% (90%, 196%)	0.84 (0.64, 1.03)	44% (32%, 56%)	0.91 (0.72, 1.10)	47% (37%, 58%)	47% (37%, 58%)
SPMS at baseline and:											
Baseline EDSS ≤3.5	58	8.2	5.9	2.69	-1.48 (-1.87, -1.09)	181% (158%, 203%)	-1.10 (-1.51, -0.70)	151% (133%, 170%)	-1.05 (-1.46, -0.64)	142% (125%, 159%)	142% (125%, 159%)
Baseline EDSS 4–5.5	182	7.8	5.1	4.55	-0.77 (-0.95, -0.59)	182% (161%, 202%)	-0.01 (-0.20, 0.19)	101% (89%, 112%)	0.13 (-0.06, 0.32)	94% (85%, 102%)	94% (85%, 102%)
Baseline EDSS ≥6	405	7.8	5.1	6	-0.65 (-0.77, -0.53)	292% (255%, 329%)	0.32 (0.21, 0.44)	76% (68%, 84%)	0.39 (0.28, 0.50)	76% (69%, 83%)	76% (69%, 83%)
Date of baseline assessment											
On or before 28 August 2003	2351	8.7	6.9	3.32	0.02 (-0.06, 0.10)	99% (94%, 103%)	0.53 (0.46, 0.61)	75% (71%, 78%)	0.58 (0.51, 0.66)	75% (72%, 79%)	75% (72%, 79%)
After 28 August 2003	2511	8.8	7.0	3.05	0.21 (0.14, 0.29)	88% (83%, 92%)	0.67 (0.60, 0.75)	69% (66%, 72%)	0.71 (0.64, 0.78)	71% (67%, 74%)	71% (67%, 74%)
Baseline at or before 31 August 2003 and:											
Baseline EDSS ≤3.5	1352	8.9	7.5	2.07	0.27 (0.16, 0.38)	87% (82%, 92%)	0.57 (0.46, 0.68)	77% (72%, 81%)	0.61 (0.50, 0.72)	77% (73%, 82%)	77% (73%, 82%)
Baseline EDSS 4–5.5	631	8.6	6.6	4.42	-0.21 (-0.34, -0.08)	119% (107%, 131%)	0.49 (0.36, 0.62)	73% (66%, 80%)	0.54 (0.41, 0.66)	74% (68%, 80%)	74% (68%, 80%)
Baseline EDSS ≥6	368	7.9	5.4	6	-0.50 (-0.64, -0.36)	244% (202%, 287%)	0.48 (0.34, 0.62)	65% (56%, 74%)	0.55 (0.41, 0.69)	67% (59%, 75%)	67% (59%, 75%)
Baseline after 31 August 2003 and:											
Baseline EDSS ≤3.5	1588	9.0	7.4	1.97	0.42 (0.32, 0.52)	81% (77%, 85%)	0.69 (0.59, 0.79)	72% (68%, 76%)	0.71 (0.61, 0.81)	73% (69%, 78%)	73% (69%, 78%)
Baseline EDSS 4–5.5	644	8.5	6.8	4.43	-0.02 (-0.16, 0.11)	102% (90%, 114%)	0.68 (0.54, 0.82)	62% (55%, 69%)	0.74 (0.60, 0.87)	64% (58%, 71%)	64% (58%, 71%)
Baseline EDSS ≥6	279	8.2	5.6	6	-0.43 (-0.58, -0.27)	219% (175%, 264%)	0.57 (0.41, 0.72)	61% (50%, 71%)	0.63 (0.48, 0.78)	63% (54%, 71%)	63% (54%, 71%)

EDSS, Expanded Disability Status Scale; MS, multiple sclerosis; RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis.

**Table 5** Other subgroup analyses (Markov model only)

Population	n	Mean years		Mean	EDSS		Utility	
		At risk	On treatment	Baseline EDSS	Absolute treatment effect	Relative rate of progression	Absolute treatment effect	Relative rate of progression
Primary analysis population	4862	8.7	7.0	3.18	0.12 (0.07, 0.17)	93% (90%, 96%)	0.039 (0.034, 0.044)	76% (73%, 79%)
One-way analyses								
On treatment throughout	2855	8.7	8.7	3.02	0.46 (0.39, 0.53)	73% (69%, 77%)	0.073 (0.067, 0.079)	55% (51%, 58%)
Ever off-treatment	2007	8.8	4.6	3.40	-0.36 (-0.44, -0.28)	123% (118%, 128%)	-0.010 (-0.018, -0.001)	106% (101%, 111%)
By gender								
Male	1224	8.6	6.8	3.19	-0.12 (-0.23, 0.00)	107% (100%, 114%)	0.013 (0.002, 0.024)	92% (85%, 99%)
Female	3638	8.8	7.0	3.17	0.20 (0.14, 0.26)	88% (84%, 92%)	0.048 (0.042, 0.053)	71% (67%, 74%)
Baseline at or before 28 February 2003								
Baseline after 28 February 2003	3865	8.8	7.0	3.13	0.14 (0.09, 0.20)	91% (88%, 95%)	0.040 (0.034, 0.045)	75% (72%, 79%)
Baseline at or before 31 August 2003								
Baseline after 31 August 2003	2511	8.8	7.0	3.05	0.21 (0.14, 0.29)	88% (83%, 92%)	0.044 (0.037, 0.052)	73% (69%, 77%)
Disease duration at baseline								
≤3 years	1251	8.8	7.3	2.46	0.45 (0.33, 0.57)	77% (71%, 83%)	0.060 (0.050, 0.070)	65% (59%, 71%)
>3 years	3611	8.7	6.9	3.43	0.00 (-0.05, 0.06)	100% (96%, 103%)	0.032 (0.026, 0.037)	80% (76%, 84%)
≤6 years	2295	8.8	7.2	2.67	0.32 (0.24, 0.40)	83% (78%, 87%)	0.051 (0.043, 0.058)	70% (65%, 74%)
>6 years	2567	8.7	6.8	3.64	-0.06 (-0.12, 0.01)	104% (99%, 108%)	0.028 (0.022, 0.035)	82% (78%, 86%)
Two-way analyses								
Disease duration ≤3 years and:								
Baseline EDSS ≤3.5	981	8.9	7.5	1.80	0.60 (0.47, 0.73)	73% (67%, 79%)	0.068 (0.058, 0.079)	61% (55%, 67%)
Baseline EDSS 4–5.5	199	8.4	6.7	4.41	0.00 (-0.29, 0.28)	100% (74%, 127%)	0.039 (0.009, 0.068)	74% (56%, 94%)
Baseline ≥EDSS 6	71	8.0	5.7	6	-0.35 (-0.76, 0.05)	209% (83%, 338%)	0.001 (-0.050, 0.051)	100% (65%, 135%)
Disease duration >3 years and:								
Baseline EDSS ≤3.5	1959	9.0	7.4	2.12	0.23 (0.14, 0.31)	90% (85%, 94%)	0.044 (0.036, 0.051)	73% (69%, 78%)
Baseline EDSS 4–5.5	1076	8.6	6.7	4.43	-0.14 (-0.23, -0.04)	112% (104%, 121%)	0.038 (0.027, 0.048)	76% (69%, 82%)
Baseline ≥EDSS 6	576	8.0	5.4	6	-0.48 (-0.59, -0.37)	236% (204%, 268%)	-0.021 (-0.038, -0.004)	114% (103%, 125%)

EDSS, Expanded Disability Status Scale.

8.7 and 10 in our study). The use of marginal structured Cox modelling (which included the contemporary untreated patients) did not change interpretation of findings.<sup>30</sup>

**Selection bias**

UK patients eligible for DMTs prior to 2002 but who progressed rapidly so they were no longer eligible would have been excluded. This bias would favour a treatment effect. Working against a treatment effect is the exclusion of those who were previously eligible but were no longer active enough to be included. A bias against a treatment effect might have occurred due to the presence of prevalent as well as incident ABN eligible patients, which would have included patients with later disease being treated,

and it is possible that DMTs may be less effective later on in the disease course. Indeed our results suggested a greater treatment effect when excluding patients recruited in the first 18 months (once the backlog of prevalent patients had been entered). It is also reassuring that disease duration at study entry was almost identical for both data sets (table 3).

Other methodological issues worth considering are the use of the EDSS as the main outcome measure and the requirements for the EDSS scoring. Although the EDSS is not sensitive to change over the shortterm and focuses on mobility, it is the most widely used and accepted measure of disability in MS. Using this measure also allows us to convert the outcome to quality of life measures for cost-effectiveness calculations and to

**Table 6** Variation of the treatment effect with year of analysis (MLM)

Analysis	Relative progression rate (%) at year:				
	2	4	6	8	10
Primary analysis cohort					
EDSS basis (95% CI)	47% (41% to 52%)	58% (54% to 62%)	66% (63 to 69%)	71% (68 to 74%)	72% (69 to 74%)
Utility basis (95% CI)	36% (31% to 41%)	55% (51% to 60%)	68% (64 to 72%)	76% (73 to 80%)	77% (74 to 80%)
RRMS-only subgroup					
EDSS basis (95% CI)	42% (36% to 48%)	55% (51% to 59%)	63% (60% to 66%)	69% (66% to 72%)	69% (67% to 72%)
Utility basis (95% CI)	28% (22% to 34%)	50% (45% to 55%)	62% (58% to 66%)	74% (70% to 78%)	76% (72% to 79%)

EDSS, Expanded Disability Status Scale; MLM, multilevel model; RRMS, relapsing remitting multiple sclerosis.



assess whether our results are plausible by comparing the results with other studies. The clinicians in our study did not undergo formal EDSS training as is required in clinical trials, rather the EDSS was collected within the routine clinical setting. Sites were instructed to use the same clinician to perform the EDSS over time where possible, and it was performed mainly by neurologists with expertise in MS. This was similar to the way the British Columbia natural history data set EDSS measurements were performed and thus comparable. The interobserver variability in our study was previously reported<sup>31</sup> to have a kappa value of 0.59, 0.71 and 0.85: for full agreement, within 0.5 and within 1.0 EDSS scores, respectively, and it is expected that the intraobserver variability is even better.

On balance, therefore, we consider that there are sound reasons for believing that these results represent a true treatment effect. These reasons include the good agreement between our estimates of the absolute treatment effect at year 2 of a mean EDSS difference of 0.22<sup>13</sup> and those derived from a meta-analysis of RCTs which gave a mean EDSS difference of 0.25<sup>32</sup>; the generally good agreement between two very different modelling approaches (particularly when the Markov models allows time to be included in the model); the adjustments included in each model to allow for the (relatively small) differences in baseline characteristics between the British Columbia and RSS data sets; and the use of a wide range of sensitivity analyses to test for possible residual sources of bias.

### Implications for the treatment of MS

Assessing the longer term effectiveness of treatment is vital in considering the cost-effectiveness of expensive treatments for chronic conditions, but RCTs have not and cannot address this in MS where disability is acquired over many years. We have performed the only prospective long-term study of the treatment effect of MS DMTs, and have managed to recruit and retain a very large cohort (with higher levels than many short-term RCTs). Our results show that the benefits seen in the short term for these drugs are maintained over a 10-year period, although the treatment effect appears to decrease over time and may not justify the prices for these drugs in some healthcare systems particularly where the drug costs are higher.<sup>33</sup> This apparent waning effect would be consistent with the recent meta-analysis<sup>34</sup> showing an inverse age-dependent association with efficacy. Because age, disease duration and EDSS scores are all inter-related and the latter two were associated with the treatment effect size in our study, it is not clear what is the driving factor. Although newer treatments are now available and are often preferred because of greater efficacy or ease of administration, their cost-effectiveness has in turn been assessed by NICE using incremental comparison with these first-generation drugs and more recently in direct head-to-head studies now that placebo studies are regarded as unethical. Thus the results of the MS RSS will have important consequences for assessing the cost-effectiveness of subsequent currently licensed and future drugs for MS.

Our finding that the treatment effect of these drugs is attenuated over time could have important implications for clinical practice, both for MS and perhaps for other longer term conditions. It might now be ethical to conduct an RCT to determine whether discontinuation of treatment—after a given period on treatment, or after reaching a given level of disability—offers any disadvantage over continuing treatment. An RCT is currently under way to assess the effects of withdrawing treatment in older patients with MS who have had no relapses or new brain lesions for at least 5 years.<sup>35</sup> Future trials of treatments for MS could be

designed so that one arm receives treatment for an initial period and then discontinues treatment. On the basis of our findings, it might be more cost-effective—and ultimately in the interests both of patients with MS and of patients with other conditions—to intervene earlier in the disease course, rather than to persist with treatment for long periods.

### Author affiliations

<sup>1</sup>Clinical Neurology, The Oxford University Hospitals Trust, Oxford, UK

<sup>2</sup>Department of Neurology, The Newcastle upon Tyne Hospitals Trust, Newcastle upon Tyne, UK

<sup>3</sup>Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

<sup>4</sup>Biostatistics, PAREXEL International, Berlin, Germany

<sup>5</sup>Department of Medicine (Neurology), University of British Columbia, Vancouver, British Columbia, Canada

<sup>6</sup>The Townsville Hospital, Townsville, Queensland, Australia

<sup>7</sup>Institute of Psychological Medicine and Clinical Neuroscience, Cardiff University, University Hospital of Wales, Cardiff, UK

<sup>8</sup>Warwick Medical School, University of Warwick, Warwick, UK

<sup>9</sup>Department of Health, Leeds, UK

**Acknowledgements** We acknowledge the input of Pelham Barton and Richard Gray, who were members of the SAG, and Nicola Russell and Tracy Nicholson, who represented the MS Trust throughout this project. The multilevel modelling project was funded by the NIHR Health Technology Assessment programme (HTA project 10/55/01) and has been published in full in the Health Technology Assessment journal series.<sup>23</sup> Visit the HTA programme website for more details (<http://www.nets.nihr.ac.uk/projects/hta/105501>). 72 centres have participated in the collection of data for the risk-sharing scheme from across the UK. Data have been collected over a 10-year period, and thus changes have been seen within the clinical teams at the centres. The authors are grateful to all the clinicians, nurses and administrators for the work undertaken at the centres, which made the RSS possible. A list of the centres can be found at <http://www.ms-trust.org.uk/research/risksharingscheme>. The MS Trust has performed an administrative coordinating role in the risk-sharing scheme, and PAREXEL has performed the data collection from the sites and undertaken the analytical work. The British Columbia MS database has been maintained over the years by grants from various sources, including the MS Society of Canada, the MS/MRI Research Group, the US National MS Society and the Canadian Institutes of Health Research. We gratefully acknowledge the BC MS Clinic neurologists who contributed to the study through patient examination and data collection (current members listed on <https://www.ms-trust.org.uk/risk-sharing-scheme-BC-clinic-neurologists>). RL and YB-S are supported by the NIHR Collaboration for Leadership in Applied Health Research and Care, West Midlands and West, respectively.

**Contributors** JP, MD and MB were the clinical leads for the study involved in the day-to-day running of the study. JP and MD were responsible for writing the paper and advising on the interpretation of the results. MB was responsible for revising the paper. ML, KT and YB-S performed the MLM, revised the paper and interpreted the results; in addition YB-S was on the scientific advisory committee. BP and TB were responsible for the Markov model analysis, revising the paper and advising on interpretation of the results. NR supplied the Welsh natural history data set and was involved in the validation of the MLM and revising the paper. HT, FZ and JO oversaw the use of their BC natural data set, revised the paper and interpreted the results. RL was the chair of the scientific advisory board, advised on the analysis and interpretation of the results, and revised the paper. CD was responsible for the oversight of the study, the analysis and interpretation of both models, and writing and revising the paper.

**Funding** The administration costs of the risk-sharing scheme were split equally five ways among the UK health departments and the four manufacturers of the DMTs (Biogen Idec, Merck Serono, Bayer, Teva UK). Additional funding from the UK Health Technology Assessment programme was obtained for the development of the MLM model (HTA Project: 10/55/01). The views and opinions expressed therein are those of the authors.

**Competing interests** None declared.

**Patient consent** Obtained.

**Ethics approval** Ethical approval for the RSS was given by the South East Medical Research Ethics Committee (MREC 2/01/78). The University of British Columbia's Clinical Research Ethics Board approved the study.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** The data storage governance is under review to assess who will store the data on behalf of the MS Trust, the four pharma companies and the Department of Health. Also the oversight committee to give permission for access to data is being discussed currently.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0>

## REFERENCES

1. Tremlett H, Zhao Y, Rieckmann P, *et al.* New perspectives in the natural history of multiple sclerosis. *Neurology* 2010;74:2004–15.
2. 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. The IFNB Multiple Sclerosis Study Group. *Neurology* 1993;43:655–61.
3. Johnson KP, Brooks BR, Cohen JA, *et al.* Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. *Neurology* 1995;45:1268–76.
4. Jacobs LD, Cookfair DL, Rudick RA, *et al.* Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann Neurol* 1996;39:285–94.
5. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group. *Lancet* 1998;352:1498–504.
6. Placebo-controlled multicentre randomised trial of interferon beta-1b in treatment of secondary progressive multiple sclerosis. European Study Group on interferon beta-1b in secondary progressive MS. *Lancet* 1998;352:1491–7.
7. Secondary Progressive Efficacy Clinical Trial of Recombinant Interferon-Beta-1a in MS (SPECTRIMS) Study Group. Randomized controlled trial of interferon-beta-1a in secondary progressive MS: Clinical results. *Neurology* 2001;56:1496–504.
8. Panitch H, Miller A, Paty D, *et al.* Interferon beta-1b in secondary progressive MS: results from a 3-year controlled study. *Neurology* 2004;63:1788–95.
9. National Institute for Clinical Excellence. *Beta interferon and glatiramer acetate for the treatment of multiple sclerosis. NICE Technology Appraisal Guidance No. 32.* London: NICE, 2002.
10. Department of Health. *Cost effective provision of disease modifying therapies for people with multiple sclerosis.* London: Health Service Circular (2002/004), Stationery Office, 2002.
11. Boggild M, Palace J, Barton P, *et al.* Multiple sclerosis risk sharing scheme: two year results of clinical cohort study with historical comparator. *BMJ* 2009;339:b4677.
12. Palace J, Bregenzer T, Tremlett H, *et al.* UK multiple sclerosis risk-sharing scheme: a new natural history dataset and an improved Markov model. *BMJ Open* 2014;4:e004073.
13. Palace J, Duddy M, Bregenzer T, *et al.* Effectiveness and cost-effectiveness of interferon beta and glatiramer acetate in the UK Multiple Sclerosis Risk Sharing Scheme at 6 years: a clinical cohort study with natural history comparator. *Lancet Neurol* 2015;14:497–505.
14. Lawton M, Tilling K, Robertson N, *et al.* A longitudinal model for disease progression was developed and applied to multiple sclerosis. *J Clin Epidemiol* 2015;68:1355–65.
15. Chilcott J, Miller D, McCabe C. Modelling the cost effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis. *BMJ* 2003;326:522.
16. Parexel/Department of Health. "Cost effective provision of disease modifying therapies for people with Multiple Sclerosis: Statistical Analysis Plan for year 10 data" (Department of Health, November 2015).
17. Association of British Neurologists "Guidelines for the use of beta interferons and glatiramer acetate in multiple sclerosis". January 2001.
18. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;33:1444–52.
19. Lechner-Scott J, Kappos L, Hofman M, *et al.* Can the Expanded Disability Status Scale be assessed by telephone? *Mult Scler* 2003;9:154–9.
20. Shirani A, Zhao Y, Karim ME, *et al.* Association between use of interferon beta and progression of disability in patients with relapsing-remitting multiple sclerosis. *JAMA* 2012;308:247–56.
21. Jackson CH, Sharples LD, Thompson SG, *et al.* Multistate Markov models for disease progression with classification error. *Journal of the Royal Statistical Society: Series D* 2003;52:193–209.
22. Goldstein H, Models MS. *Multilevel Statistical Models (second edition).* London: Edward Arnold, 1995.
23. Tilling K, Sterne JA, Wolfe CD. Multilevel growth curve models with covariate effects: application to recovery after stroke. *Stat Med* 2001;20:685–704.
24. Tilling K, Lawton M, Robertson N, *et al.* Modelling disease progression in relapsing-remitting onset multiple sclerosis using multilevel models applied to longitudinal data from two natural history cohorts and one treated cohort. *Health Technol Assess* 2016;20:1–48.
25. Gold R, Kappos L, Arnold DL, *et al.* Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N Engl J Med* 2012;367:1098–107.
26. Fox RJ, Miller DH, Phillips JT, *et al.* Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *N Engl J Med* 2012;367:1087–97.
27. Lourenco P, Shirani A, Saeedi J, *et al.* Oligoclonal bands and cerebrospinal fluid markers in multiple sclerosis: associations with disease course and progression. *Mult Scler* 2013;19:577–84.
28. Inusah S, Sormani MP, Cofield SS, *et al.* Assessing changes in relapse rates in multiple sclerosis. *Mult Scler* 2010;16:1414–21.
29. Shirani A, Zhao Y, Kingwell E, *et al.* Temporal trends of disability progression in multiple sclerosis: findings from British Columbia, Canada (1975–2009). *Mult Scler* 2012;18:442–50.
30. Karim ME, Gustafson P, Petkau J, *et al.* Marginal structural Cox models for estimating the association between  $\beta$ -interferon exposure and disease progression in a multiple sclerosis cohort. *Am J Epidemiol* 2014;180:160–71.
31. Pickin M, Cooper CL, Chater T, *et al.* The Multiple Sclerosis Risk Sharing Scheme Monitoring Study – early results and lessons for the future. *BMC Neurol* 2009;9:1.
32. Filippini G, Munari L, Incurvaia B, *et al.* Interferons in relapsing remitting multiple sclerosis: a systematic review. *The Lancet* 2003;361:545–52.
33. Hartung DM, Bourdette DN, Ahmed SM, *et al.* The cost of multiple sclerosis drugs in the US and the pharmaceutical industry: Too big to fail? *Neurology* 2015;84:2185–92.
34. Weideman AM, Tapia-Maltsos MA, Johnson K, *et al.* Meta-analysis of the age-dependent efficacy of multiple sclerosis treatments. *Front Neurol* 2017;8:577.
35. Explore Our Portfolio of Funded Projects. [http://www.pcori.org/research-results/2015/discontinuation-disease-modifying-therapies-dmts-multiple-sclerosis-ms](http://www.pcori.org/research-results/2015/discontinuation-disease-modifying-therapies-dmts-multiple-sclerosis-mshttp://www.pcori.org/research-results/2015/discontinuation-disease-modifying-therapies-dmts-multiple-sclerosis-ms)