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

Kate Tilling,[1]* Michael Lawton,[1] Neil Robertson,[2]
Helen Tremlett,[3] Feng Zhu,[3] Katharine Harding,[2]
Joel Oger[3] and Yoav Ben-Shlomo[1]

[1]School of Social and Community Medicine, Bristol University, Bristol, UK
[2]Department of Neurology, Institute of Psychological Medicine and Clinical Neuroscience, Cardiff University, Cardiff, UK
[3]Faculty of Medicine, Department of Medicine, Division of Neurology, University of British Columbia, Vancouver, BC, Canada

*Corresponding author

Declared competing interests of authors: Kate Tilling has received non-financial support from the Multiple Sclerosis (MS) Trust during the conduct of the study, and had her expenses paid by the MS Trust to attend meetings of the UK MS risk-sharing scheme (RSS) scientific advisory group in order to outline the plan for these analyses. Michael Lawton has had his expenses paid by the MS Trust to attend meetings of the UK MS RSS scientific advisory group in order to outline the plan for these analyses, and also his travel and accommodation expenses for visiting Vancouver to analyse the British Columbia MS data set. Neil Robertson has received travel grants and honoraria from Biogen, Novartis, Serono, Sanofi, Genzyme and Bayer and holds grants for unrelated work from Genzyme and Novartis. Helen Tremlett is funded by the MS Society of Canada (Don Paty Career Development Award), Michael Smith Foundation for Health Research and is the Canada Research Chair for Neuroepidemiology and MS. She has received research support from the National MS Society, the Canadian Institutes of Health Research and the UK MS Trust; and speaker honoraria and/or travel expenses to attend conferences from the Consortium of MS Centres (2013), the National MS Society (2012, 2014), Bayer Pharmaceuticals (2010), Teva Pharmaceuticals (2011), European Committee for Treatment and Research in MS (2011, 2012, 2013 and 2014), UK MS Trust (2011), the Chesapeake Health Education Program, US Veterans Affairs (2012), Novartis Canada (2012), Biogen Idec (2014) and the American Academy of Neurology (2013, 2014 and 2015). Unless otherwise stated, all speaker honoraria are either donated to a MS charity or to an unrestricted grant for use by her research group. Yoav Ben-Shlomo has had his expenses paid by the MS Trust to attend meetings of the UK MS RSS Scientific Advisory Group in order to outline the plan for these analyses, and has a relative with MS who is currently on treatment for the disease. He is a member of the UK MS RSS Scientific Advisory Group.

Published October 2016
DOI: 10.3310/hta20810

Scientific summary

Predicting disease progression in patients with MS

Health Technology Assessment 2016; Vol. 20: No. 81
DOI: 10.3310/hta20810

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Multiple sclerosis (MS) is a chronic inflammatory and neurodegenerative disorder in which the clinical features, including presentation, disease course and rates of accumulation of neurological disability, demonstrate high degrees of individual variation. The majority of patients (≈85%) will present with a relapsing–remitting (RR) disease course, typified by acute relapses ('attacks') followed by periods of remission (recovery). Over time, relapsing–remitting multiple sclerosis (RRMS) can convert to secondary-progressive multiple sclerosis (SPMS), which is associated with a progressive course (gradual worsening) with or without superimposed relapses. The variability in the MS disease course can cause anxiety and frustration for patients, families and health-care professionals alike. Better prediction of disease progression represents a major unmet need in MS. A realistic prognosis would be of considerable importance to patients and would also help to inform therapeutic and management choices. This is particularly relevant when recent developments in therapeutics are hoped to have a significant positive impact on the natural history of MS, but there are also clearly costs in terms of serious adverse events. Patients and their carers increasingly request access to relevant prognostic information to help plan both professional and family life decisions.

There is no cure for MS, but available treatment includes disease-modifying therapies (DMTs), also known as 'immunomodulatory agents', which are currently licensed only for patients with relapsing-onset MS (RR or SPMS). In 2002, there were four licensed DMTs: two forms of interferon beta-1a [Avonex® (Biogen Idec Ltd) and Rebif® (Merck Serono Ltd), with high- and low-dose formulations available, 44 µg and 22 µg, respectively], one form of interferon beta-1b [Betaferon® (Bayer Plc)] and glatiramer acetate [Copaxone® (Teva UK Ltd)] (first generation). At that time, the UK's National Institute for Health and Care Excellence appraised the evidence for cost-effectiveness of these DMTs and concluded that they were not cost-effective over a 10- or 15-year period. However, they also concluded that, because the relevant clinical trials were based on only short-term outcomes, there was not enough information to assess the longer-term implications of these DMTs. To try to address these concerns, the UK MS risk-sharing scheme (RSS) was established in 2002. Under the scheme, the UK's Department of Health agreed to pay for the DMTs, conditional on treated patients being included in a 10-year monitoring study aiming to assess progression of the disease. The monitoring study aimed to compare progression of MS within this treated group of patients with progression in an untreated cohort. Initially, a cohort of untreated patients from London, Ontario, Canada, were selected [being a subset of 314 patients judged to have fulfilled the Association of British Neurologists (ABN) criteria, from the London, Ontario, cohort of 1043 patients recruited between 1972 and 1984]. The first report from the UK MS RSS based on the 2-year follow-up data was unable to find evidence for their cost-effectiveness, but the authors were hesitant in drawing any firm conclusions because of methodological limitations. The main concerns raised related to the use of the London, Ontario, cohort, which included only patients with stable or deteriorated disability scores; in contrast, in real-life clinical practice, scores naturally fluctuate and can also improve. This use of the London, Ontario, cohort forced a complex algorithm to be used in the analysis, creating an artificial 'no-improvement' rule. This caused multiple problems, resulting in uncertainty in the findings, with the sensitivity analysis revealing qualitatively different results from the main analysis (suggesting that the DMTs might actually have a beneficial effect). The authors noted that, given potential secular trends in the natural history of MS, a more contemporary cohort would be preferable, and that modelling approaches other than the Markov model could be used.

We set out to develop and validate models that would facilitate predicting disease progression in patients with RR-onset MS. We then aimed to use these models to estimate the impact of the current first generation of disease-modifying drugs on disease progression in MS.

© Queen's Printer and Controller of HMSO 2016. This work was produced by Tilling *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Objectives

1. To develop and apply multilevel models to repeated measures (longitudinal) data on disease progression in patients with relapsing–remitting onset MS (including those who go on to reach SPMS) in two different countries with similar health systems, by accessing the University of Wales Multiple Sclerosis (UoWMS) and British Columbia Multiple Sclerosis (BCMS) databases.
2. To use these models to determine average progression trajectory [as measured by the Expanded Disability Status Scale (EDSS)], individual deviations from the average and accuracy of prediction of individual trajectories.
3. To examine the impact of 'informative censoring' using joint modelling of the EDSS trajectory and time to censoring (whether censored because a patient drops out or because they cease to contribute to the analysis when they start a DMT) and, if necessary, use multiple imputation methods as part of a sensitivity analysis.
4. To apply the multilevel model developed above to data from the 6-year follow-up of the UK MS RSS cohort to estimate average progression trajectory, individual deviations from the average and average difference between individual progression trajectories in the UK MS RSS and progression predicted by the BCMS model and, thus, to estimate the effect of the MS drugs on progression.
5. To apply patient-derived utilities to the observed EDSS score in the UK MS RSS cohort and to the EDSS score predicted under the natural history model, and thus to estimate the difference in utility that may be a result of DMT. Utility is a measure of society's perception of the quality of life of a patient in a given state of health. A utility of 1 represents perfect health while a utility of 0.5 implies that, on average, people would regard 12 months of life in that health state as equally preferable to 6 months of life in perfect health.

Methods

We used the Welsh database (UoWMS) for exploratory modelling, then developed the final natural history model using the BCMS database and validated its ability to predict progression in UoWMS patients. Patients with RRMS or SPMS were included in the models once they had met the ABN criteria for eligibility for starting DMT. We operationalised these criteria as age ≥ 18 years, EDSS score ≤ 6.5 points and occurrence of at least two relapses in the previous 2 years. From the BCMS cohort we included only patients who reached ABN eligibility between 1980 and 1997, who had definite MS and a minimum of two EDSS scores at least 9 months apart. We excluded EDSS scores taken during a relapse, or when disability was affected by other factors considered largely unrelated to MS (e.g. hip fracture), or once a patient started immunomodulatory drug treatment. Multilevel models were used to model the trajectory of EDSS scores with time, and to take account of the fact that observations of EDSS score (level 1) are nested within individuals (level 2), and that EDSS score observations within the same individual are likely to be correlated. These models allow EDSS trajectory to vary between individuals, and observed EDSS score to vary about this trajectory within individuals. The model developed for the natural history of MS (using the BCMS cohort) was used to predict EDSS score for patients in the UoWMS cohort. The final model (developed using the BCMS cohort) was applied to the UK MS RSS cohort, and used to estimate the EDSS score had the cohort not been treated. This predicted 'natural history' EDSS score was compared with the observed EDSS score at 6 years' follow-up in this treated cohort, to estimate the effect of treatment.

Results

First, we were able to show model replication between the two geographically distinct cohorts, which is remarkable in such a variable disease as MS (aims 1 and 2). Whether or not the time axis was time since onset or time since ABN eligibility, the Welsh and British Columbia cohorts gave rise to models with the same functional form and similar parameters. For example, when modelling EDSS score against time since onset, the models from the UoWMS and BCMS cohorts showed a significant difference only in the level 1 constant random effect and the constant fixed effect. A difference in the constant fixed effect was expected, given that, from the patients' characteristics, the first 'ABN eligible' EDSS score was, on average, higher in the UoWMS cohort. The difference in the level 1 constant random effect shows that there was less variation around the initial EDSS score for the UoWMS cohort than the BCMS cohort.

The coefficients from the model for one cohort could be used to predict EDSS score for patients from the other cohort by conditioning on the first EDSS score observed for an individual. The accuracy of these predictions will always be limited by measurement error associated with the EDSS score. Thus, when using the BCMS model to predict EDSS score for individual patients in the UoWMS cohort, conditional on their first EDSS score observation, 90% of the predictions lay within –2.84 to 1.56 points of the true EDSS score value. Although the accuracy of these predictions may not be sufficiently good for predicting at an individual level for clinical purposes, these results provide strong evidence of the ability to use one cohort of MS patients as a non-randomised comparison group for another.

We confirmed several factors as being related either to average EDSS score or to rate of change of EDSS score over time. Those who were older at MS onset tended to have higher EDSS scores and slightly greater rates of increase in EDSS score over time. Those with SPMS had a higher EDSS score at ABN eligibility (although no obvious difference in progression), as did those with more relapses in the 2 years prior to ABN eligibility. Women had lower EDSS scores, on average, than men, but again with no obvious differences in EDSS trajectory. The estimates of these associations were consistent between the two cohorts.

Our study indicates that EDSS trajectory can be modelled using multilevel models. In all models, there was no strong evidence of non-normality of residuals and, therefore, no need to consider further transformations of EDSS score. We also found no evidence against the use of the EDSS score as a continuous variable, rather than an ordinal outcome (as it was designed), in that the models with EDSS score as a continuous outcome appeared to predict EDSS score well. There was evidence that intraindividual variability in EDSS score was greater for lower EDSS scores, as previously suggested.

Using the BCMS cohort natural history model to predict what EDSS score would have been had a treated cohort (the UK MS RSS) not been treated leads to the conclusion that the UK MS RSS may have a slightly lower mean EDSS score [by 0.59 points, 95% confidence interval (CI) 0.54 to 0.64 points] at 6 years post treatment (aim 3). Using patient-derived utility scores (on a scale of 0 to 1), the UK MS RSS showed a slightly higher utility score at 6 years' follow-up than that predicted using the natural history cohorts (by 0.04 utility points, 95% CI 0.036 to 0.047 utility points) (aim 4).

Conclusions

Our major finding was the ability to develop and then validate a model of disease progression in two geographically distinct MS cohorts (aim 1). Our trajectory of disability (EDSS score) over time was similar in the Welsh and BCMS cohorts, with no evidence of informative censoring (aim 2). This is a significant finding given the high variability of progression in MS and the often-differing findings from the MS natural history studies published to date. This provides evidence that using appropriately collected disability data (EDSS score) and advanced modelling, the progression of MS can be modelled, and that models from one population can be used to provide information about a different population of MS patients. Using this approach, we concluded that EDSS score and EDSS score-based utility progression in the treated UK MS RSS cohort was slower than that predicted by the natural history model (aims 3 and 4). Further research into the ability of these models to provide useful information to individuals (rather than populations) is needed before they can be used to provide prognostic information and help to decide on treatment options.

© Queen's Printer and Controller of HMSO 2016. This work was produced by Tilling *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Future research should focus on long-term follow-up of treated and untreated cohorts and randomised controlled trials, in order to examine the robustness of these results, and the extent to which they may be causal.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Health Technology Assessment HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

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.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 10/55/01. The contractual start date was in August 2011. The draft report began editorial review in August 2015 and was accepted for publication in May 2016. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2016. This work was produced by Tilling *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Health Technology Assessment Editor-in-Chief

Professor Hywel Williams Director, HTA Programme, UK and Foundation Professor and Co-Director of the Centre of Evidence-Based Dermatology, University of Nottingham, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

Professor John Norrie Health Services Research Unit, University of Aberdeen, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk