Ocrelizumab for treating patients with primary progressive multiple sclerosis: An Evidence Review Group perspective of a NICE Single Technology Appraisal

Short running header: Ocrelizumab for primary progressive multiple sclerosis: an ERG perspective

Peter Auguste[1], Jill Colquitt[2], Martin Connock[1], Emma Loveman[2], Rachel Court[1], Olga Ciccarelli[3], Carl Counsell[4], Xavier Armoiry[1,5]

[1] Warwick Evidence, Warwick Medical School, University of Warwick, Coventry, CV4 7AL, UK

[2] Effective Evidence, UK

[3] University College London (UCL) Queen Square Institute of Neurology, UK

[4] Institute of Applied Health Sciences, University of Aberdeen, UK

[5] Lyon University, Claude Bernard University Lyon (UMR CNRS MATEIS), School of Pharmacy (ISPB)/ Edouard Herriot Hospital, Pharmacy department, 69008 Lyon, France

Peter Auguste	P.Auguste@warwick.ac.uk
Jill Colquitt	Jill.colquitt@effectiveevidence.org
Martin Connock	M.Connock@warwick.ac.uk
Emma Loveman	Emma.loveman@effectiveevidence.org
Rachel Court	R.A.Court@warwick.ac.uk
Olga Ciccarelli	O.ciccarelli@ucl.ac.uk
Carl Counsell	Carl.counsell@abdn.ac.uk
Xavier Armoiry	xavier.armoiry@univ-lyon1.fr

*Corresponding author:
Peter Auguste
Warwick Medical School
University of Warwick
Gibbet Hill Road
Coventry, UK, CV4 7AL
Tel: +44 (0)2476 150648
Email: P.Auguste@warwick.ac.uk

Abstract

Ocrelizumab is indicated for relapsing remitting and primary progressive multiple sclerosis (RRMS and PPMS, respectively). In an appraisal undertaken by the National Institute of Health and Care Excellence, the company Roche, presented the evidence for ocrelizumab used in patients with PPMS which came from one single randomised controlled trial (RCT), comparing ocrelizumab versus placebo. Based on results from this trial, the licensed indication was restricted to patients with early PPMS in terms of disease duration and level of disability, and with imaging features characteristic of inflammatory activity. Overall, the Evidence review group (ERG) found that the RCT had a low risk of bias. In the post-hoc defined magnetic resonance imaging (MRI) active subgroup, matching the label indication, the risk of confirmed disability progression sustained for 12 weeks (CDP-12) was significantly delayed in the ocrelizumab group compared to placebo. However, considering the same risk with progression sustained for 24 weeks (CDP-24), which was deemed the most clinically relevant, the benefit from ocrelizumab did not reach statistical significance. In the same MRI active subgroup, benefits from ocrelizumab on functional outcomes and on health-related quality of life were not clearly demonstrated.

A *de novo* Markov model was used to estimate the cost-effectiveness of ocrelizumab versus best supportive care (BSC) for treating patients with PPMS. Health states were defined by expanded disability status scale (EDSS) ranging from 0-9. Disability progression was based on the MSBase natural history cohort which exhibited disease progression in the absence of disease modifying therapy. Treatment with ocrelizumab delayed disability progression, with evidence of its clinical effectiveness obtained from the RCT. The economic analysis was undertaken from the National Health Service and Personal Social Services perspective, and the outcomes were reported in terms of life-years gained and quality-adjusted life years (QALYs), with the overall results reported in terms of an incremental cost-effectiveness ratio (ICER), expressed as cost per QALY gained over a 50-year time horizon. Both costs and effects were discounted at 3.5% per annum. The company undertook deterministic one-way sensitivity analyses and scenario analyses, including probabilistic sensitivity analysis (PSA). The ERG raised several concerns that were discussed at the appraisal committee meetings, which resulted the committee's preferences and a revised company's economic analysis.

Under an approved patient access scheme with appraisal committee preferences applied, analyses yielded an ICER of approximately £78,300 per QALY. Sensitivity analysis results indicated that the treatment effect on CDP-12 had the greatest impact. Results for the PSA showed that at a willingness-to-pay threshold of £30,000 per QALY gained, ocrelizumab versus BSC had a zero probability of being cost-effective.

Following new analyses submitted by the company, with a revised confidential patient access scheme, NICE recommended ocrelizumab in the treatment of early PPMS in adults with imaging features characteristic of inflammatory activity.

Abstract word count: 455

Total word count: 5869

Key points for decision makers:

- Clinical evidence for ocrelizumab in primary progressive multiple sclerosis (PPMS) came from a randomised controlled trial (RCT). From this trial, the European Medical Agency label indication was restricted to a post-hoc defined population subgroup with early PPMS in terms of disease duration and level of disability, and with imaging features characteristic of inflammatory activity.
- Trial results for the magnetic resonance imaging (MRI) active subgroup showed ocrelizumab benefit in confirmed disability progression (CDP) at 12 weeks, while at the clinically more relevant time point of 24-weeks benefit was not clearly demonstrated. The benefit of ocrelizumab on exploratory functional outcomes and on fatigue were unclear.
- PPMS was estimated from progression through the expanded disability status scale (EDSS), which is heavily dependent on walking ability, and fails to adequately assess upper limb function and cognitive impairment; other measures submitted by Roche (fatigue and upper limb impairment) were judged insufficiently robust to be incorporated into cost-effectiveness in the form of disutilities.
- It is important that there is transparency in the choice of outcomes to be incorporated as measures of disutility.
- National Institute for Health and Care Excellence (NICE) has recommended ocrelizumab, within its marketing authorisation, as an option for treating early PPMS in adults with imaging features characteristic of inflammatory activity.

1. Introduction

In England, the appraisal of targeted newly licensed medicines is undertaken independently by the National Institute for Health Care and Excellence (NICE) which is responsible for issuing guidance on the use of drugs within the National Health Service (NHS). Over the last decade, many new pharmaceutical products have been assessed by NICE as part of the single technology appraisals (STA) programme. In each of these appraisals, the company of a particular technology is invited to provide a company submission (CS) presenting evidence on the clinical and cost-effective use of its product for a single indication. Following this, an Evidence Review Group (ERG), in this case Warwick Evidence (the University of Warwick, UK), is commissioned by the UK Health Technology Assessment (HTA) programme to critique the clinical and economic evidence submitted by the company. This paper summarises the ERG's critique of Roche's submission to NICE pertaining to the use of ocrelizumab for treating patients living with primary progressive multiple sclerosis (PPMS) [ID 938], and provides a brief summary of the development of the NICE guidance.[1]

2. ERG critique of the Decision Problem defined by the company

Multiple sclerosis (MS) is a progressive, degenerative disease affecting the central nervous system. It is characterised by inflammation, demyelination and degeneration of neurons, and is mediated by a complex interplay between the immune system (T cells as well as B cells, antibodies, and cells of the innate immune system), glial cells (myelin-making oligodendrocytes and their precursors, microglia, and astrocytes), and neurons.[2] Unlike relapsing remitting multiple sclerosis (RRMS), in which patients experience attacks of symptoms followed by periods of improvement or remission, PPMS is typically characterised by progressive decline from onset, with occasional temporary plateaus or minor improvement.[3] PPMS has also an older age of onset, with greater susceptibility in men. [4] In the UK, PPMS represents around 14% of cases of MS.[5]

While a significant number of disease modifying therapies (DMTs) are licensed and recommended in the UK for RRMS, prior to ocrelizumab there was no active treatment licensed for PPMS and no other agents are used off-label in the UK. At the time of appraisal, established clinical management of patients with PPMS included pharmacological options to alleviate some symptoms, such as fatigue or spasticity. [6]

Compared with the population described in the NICE scope, namely patients with PPMS, the population defined in the decision problem of the CS was restricted to patients with early PPMS in terms of disease duration and level of disability, and with imaging features characteristic of inflammatory activity. This

was done for consistency with the label indication of the marketing authorisation that was granted by the European Medicine Agency (EMA) for ocrelizumab in January 2018. [7]

The ERG found the criteria of "early primary progressive multiple sclerosis (PPMS) disease in terms of disease duration and level of disability" and "with imaging features characteristic of inflammatory activity" to lack clarity and be susceptible to multiple interpretations.[7] In the absence of more precise eligibility criteria for ocrelizumab, it was felt these criteria would be at risk of being applied differently across centres, thereby creating inequalities in the access to the treatment within the NHS.

The ERG noted that the company defined inflammatory activity according to the presence of T1 gadolinium (Gd) enhancing lesions and/or active (new or enlarging) T2 lesions. Applying the criteria of new or enlarging T2 lesions to assess eligibility for ocrelizumab treatment would involve repeated magnetic resonance imaging (MRI), which was not common practice in the UK for patients with PPMS at the time of the appraisal. Similarly, many centres did not routinely undertake brain imaging with Gd. This again suggests that eligibility criteria for ocrelizumab treatment can be applied in the NHS only if there is a change in clinical practice, which, in turn, may not happen uniformly across the NHS.

Ocrelizumab is an intravenously administered medication that has been authorised for use in two indications, namely PPMS and RRMS with active disease defined by clinical or imaging features. At the time of appraisal of ocrelizumab for PPMS, NICE was also separately appraising ocrelizumab for RRMS (ID937). The latter appraisal led to ocrelizumab being recommended as an option for RRMS if alemtuzumab is contraindicated or otherwise unsuitable.[8]

Ocrelizumab is a monoclonal antibody that selectively binds to and depletes CD20-expressing B cells. The summary of product characteristics (SPC) of ocrelizumab indicates that the exact mechanisms by which it exerts its clinical effects in MS are not fully known, [7] but it is thought that it depletes CD20-expressing B cells. While the mechanism of action of ocrelizumab used in MS is novel compared to other licensed DMTs in MS (irrespective of the type of MS), the ERG considered that the selective binding to CD20+ B cells is not innovative. Indeed, the ERG noted the existence of rituximab, another anti-CD20+ monoclonal antibody currently licensed for the treatment of some hematologic malignancies and specific autoimmune disorders. Ocrelizumab and rituximab are owned by the same company, Roche Products limited. The effectiveness of rituximab in PPMS was tested in the OLYMPUS trial and no significant difference was found for time to 12-week confirmed disability progression (CDP) between rituximab and placebo after 96 weeks of follow-up (p=0.1442) [9]. Interestingly, the proportion of patients with CDP at week 96 with

rituximab was very similar to that with ocrelizumab in the ORATORIO trial [10] at week 120 (30.2% vs 32.9%, respectively). Also, subgroup analyses from the OLYMPUS trial suggested some benefit in younger patients with evidence of increased inflammatory activity, which provided some rationale to restrict inclusion to patients below the age of 55 years in the ORATORIO trial.

3. Company's original submission and outcome following appraisal committee.

3.1. Submitted Clinical Evidence and ERG critique

The main clinical evidence came from the ORATORIO (WA25046) trial, a phase III double-blind, multi-centre, placebo controlled RCT sponsored by the company. At the European Union (EU) level, the results were reviewed by the European Medicines Agency (EMA) as part of the process aimed to grant marketing authorisation. Publicly available data from ORATORIO came mainly from a paper by Montalban et al. [10].

PPMS patients were randomly assigned to either intravenous infusion of ocrelizumab 600mg (by infusion) or intravenous infusion of placebo in a 2:1 ratio for a double-blind controlled period of at least 120 weeks. Recruitment was between 3 March 2011 and 27 December 2012 and the clinical cut-off was 24 July 2015. The key inclusion criteria were age 18-55 years, PPMS diagnosis by 2005 revised McDonald criteria, expanded disability status scale (EDSS) score of 3.0-6.5 at screening, a duration of MS symptoms <15 years if EDSS score >5.0 at screening or <10 years if EDSS score ≤5.0 at screening. EDSS is a scale commonly used to measure disability progression occurring during the course of disease.

Key exclusion criteria included a history of RRMS, secondary progressive MS (SPMS) or progressive relapsing MS (PRMS), contraindications to MRI, and previous treatment with B-cell–targeted therapies and other medications for MS.

Follow-up visits occurred every 12 weeks from the date of last visit until 48 weeks had elapsed since the last treatment. Outcomes reported were confirmed disability progression (CDP) at 12 weeks (CDP-12), which was the primary outcome, CDP at 24 weeks (CDP-24), time 25-foot walk (T25FW), change in T2 brain lesions on MRI, percentage change in total brain volume, change in physical component summary score of the SF-36 (all secondary outcomes), as well as a number of exploratory endpoints, such as sustained increase in 9-hole peg test (9-HPT), and a measure of fatigue.

The main results reported: 1) from the intention-to-treat (ITT) population (which does not fully meet the licensed indication); and 2) from the post-hoc MRI active subgroup, that was defined according to active disease on MRI (presence of T1 Gd-enhancing or new T2 lesions between screening and baseline), which the company considered to be a good match to the marketing authorisation of 'early and active' disease.

Overall, the ERG found that the ORATORIO RCT had a low risk of bias. The main study quality issues discussed by the ERG were the presence of some imbalance between groups in withdrawals from the trial, which may include a perceived lack of efficacy, although the number withdrawing for this reason was not provided. It was noted that a small proportion of patients were unblinded during the trial, but the impact of this is unclear.

While the primary outcome of the trial was CDP sustained for 12 weeks, the ERG noted that the secondary endpoint CDP sustained for 24 weeks was a more clinically meaningful measure for a chronic disease like PPMS. The ERG considered that several of the secondary or exploratory endpoints measured in the trial had well known limitations. For example, the 9-HPT is the gold standard for upper limb function in MS, but it remains unclear to what extent this score contributes to detection of overall progression, and the magnitude of clinically meaningful change that can be used for the interpretation of changes in this test has been historically set at 20%.[10] 20% change in test score is often used to define clinically meaningful worsening in clinical trials involving progressive MS,[11] although the definition of this threshold needs further validation across all stages of the disease.[12] While the 9-HPT correlates well with the manual manipulation of everyday use of objects [12], the ERG noted that the use of the 9-HPT to test the ability of upper limbs to do meaningful tasks which would cause loss of independence (for example, feeding and dressing) was still unclear.

With regards to the submission, the ERG considered that outcomes were selectively reported, placing greater emphasis on statistically significant exploratory outcomes in the main submission. Several pre-defined exploratory outcomes measured in the ORATORIO trial were not presented in the main submission or its appendices.

There were no meaningful differences at baseline in demographic or disease characteristics between ocrelizumab or placebo groups in the ITT population. Similarly, there appeared to be no major imbalances on baseline characteristics for MRI active subgroup.

3.1.1. ITT population

Pre-specified analyses on the ITT population (which does not fully meet the licensed indication) showed the risk of CDP was significantly delayed in the ocrelizumab group compared to the placebo group, irrespective of whether CDP was sustained for 12 weeks (primary endpoint) (hazard ratio [HR], 0.76; 95% confidence interval [CI]: 0.59, 0.98) or 24 weeks (HR 0.75; 95% CI: 0.58, 0.98).

The benefit of ocrelizumab on health-related quality of life (HRQoL) was not consistent across the different components of the measures which were used. Indeed, there was no statistically significant

difference between ocrelizumab and placebo using the physical component score of the SF36 (SF-36 PCS) (pre-specified secondary endpoint), but there was a statistically significant improvement using the mental component score of the SF36 (SF-36 MCS) score with ocrelizumab versus placebo (exploratory endpoint). Change in EuroQoL 5 dimensions (EQ-5D) (listed as an exploratory endpoint) was not reported.

The benefit of ocrelizumab on functional outcomes (all exploratory) was unclear. While there was a statistically significant impact of ocrelizumab over placebo on upper limb function measured with the proportion of patients with ≥20% increase of the 9-HPT sustained for 12 weeks, there was no statistically significant difference between ocrelizumab and placebo in the mean change from baseline on the Multiple Sclerosis Functional Composite Score (MSFC), which measures the leg function/ambulation (using T25FW), arm/hand function (using the 9-HPT), and cognitive function (i.e., auditory information processing speed and calculation) (using the 3-seconds paced Auditory Serial Addition Test [PASAT]). Lastly, there was no statistically significant difference in the change from baseline to week 120 in the PASAT score.

3.1.2. Subgroup population matching marketing authorisation criteria

In the more relevant post-hoc MRI active subgroup, the risk of disability progression, with progression sustained for 12 or 24 weeks, was delayed in the ocrelizumab group compared to the placebo group: for 12-week CDP, the advantage reached statistical significance (HR for 12-week CDP, 0.68; 95% CI: 0.46, 0.99) while with the more relevant 24-week CDP end point it did not (HR for 24-week CDP, 0.71; 95% CI: 0.47, 1.06).

The change in T25FW from baseline to week 120 was not reported in the CS and the relative effect in reducing progression in T25FW is not known. No results for HRQoL were presented.

The benefit of ocrelizumab on functional outcomes (all exploratory) was unclear. There was a positive impact of ocrelizumab over placebo for the risk of 20% increase in 9-HPT: the HR (sustained for 12 weeks) was 0.52 (95% CI: 0.32, 0.85). However, no results on the MSFC were reported. Similarly, no results measuring the PASAT score were reported. Data provided by the company suggest that ocrelizumab had no impact on fatigue compared to placebo based on the mean changes on the modified fatigue index (MFIS).

Main clinical effectiveness results are summarised in table 1.

Given the absence of any other licensed drug for PPMS, which means no other active treatment was available, there was no evidence to support an investigation by indirect comparison of ocrelizumab versus established clinical management without ocrelizumab.

Regarding safety, the rates of events appeared to be similar between ocrelizumab and placebo in general. Treatment discontinuations due to adverse events were experienced by 4.1% in the ocrelizumab arm and 3.3% in the placebo arm. Any adverse events were experienced in 95.1% of patients in the ocrelizumab arm and 90% in the placebo arm and any serious adverse events (SAEs) by 20.4% and 22.2% for the two groups respectively. Rates of death were low in both groups. Adverse events of special interest include infusion-related reactions, infections, malignancies and anti-drug antibodies. A higher proportion of patients treated with ocrelizumab reported infusion-related reactions than placebo (39.9% ocrelizumab versus 25.5% placebo reported at least one). Overall the proportion reporting an infection was similar between groups (69.8% versus 67.8% in the ocrelizumab versus placebo groups respectively). The ERG checked and agreed that there are no other specific adverse events with a difference >3% between arms.

3.1.3.Exploratory analyses

As in the submission to the EMA, the company undertook survival analyses to evaluate the time to reach EDSS of ≥7. EDSS 7 is considered clinically important as this is when patients become essentially restricted to a wheelchair. The ERG considered that time to sustained EDSS ≥ 7 represented a more interpretable and tangible "treatment effect" than CDP 12 or 24, since this final EDSS is specified, and reaching wheelchair status would appear to be less susceptible to within- and inter-rater variability than most other EDSS transitions. Using a Weibull model, the company estimated a delay of 8.8 years in the median time to reach the EDSS 7 milestone in the ITT population. [7]

The ERG's exploratory analyses showed that for the ITT population, the gain delivered by ocrelizumab on time to progress to EDSS ≥ 7 depended heavily on the parametric models used to extrapolate beyond the observed data. Using a Weibull model (as in the submission), the delay in median time to EDSS ≥ 7 is 8.64 years in favour of ocrelizumab compared to placebo, while using a Gompertz model, the delay in median time to EDSS ≥ 7 falls to 3.06 years in favour of ocrelizumab compared to placebo (Figure 1). There was little to choose between models on the basis of information criteria; Gompertz models gave a slightly better visual fit to the KM plots. The placebo Gompertz model (whole population) conforms in shape to MSBase data (i.e. an initial increase in slope followed by decreasing slope), whereas for the Weibull models the slope of the extrapolation continuously decreases into the future. On balance, the ERG favoured the Gompertz model which predicts substantially less delay in progression of disability than the

Weibull models. Similarly, exploratory analyses of the MRI active subgroup estimated delay in median time to progressing to EDSS ≥ 7 of 2.88 years in favour of ocrelizumab using a Gompertz model, and 9.24 years using a Weibull model (Figure 1).

3.2. Submitted Cost-effectiveness evidence and ERG critique

The company undertook an extensive bibliographic search to identify studies that assessed the cost-effectiveness of DMTs for the management of patients with PPMS. This did not identify any published studies; however, their website search identified a report of an economic analysis that compared ocrelizumab with best supportive care in patients with MS, but the results were not presented in the form of an incremental cost-effectiveness ratio because the list price of ocrelizumab was not available at the time of analysis.[13] The search also sought to identify studies that reported resource use and cost information, and HRQoL for patients living with PPMS.

The company built a *de novo* Markov model programmed in Microsoft Excel to depict the natural history of patients with PPMS. The natural history of PPMS was characterised by a series of progressive health states representing the increasing levels of disability resulting from progressive loss of neurological function. The Kurtzke EDSS [14] was used to measure neurological disability and its progression overtime, with EDSS health states ranging from 0-9 and dead. Disability progression was based on patients with PPMS from the MSBase registry,[15] which showed disease progression in the absence of DMTs. The model simulated patients' disposition over the model time horizon, starting with a baseline distribution reflecting the ORATORIO trial population at recruitment. The treatment effect in the form of a hazard ratio (for CDP-12) was applied to the forward transitions, but not backward transitions. Annual cycles were used to show the movement of patients through the model. In each cycle, patients remained in these health states, after which they could progress to more severe EDSS health states or death. It was assumed that patients who discontinued treatment (due to adverse events or progressing to EDSS ≥ 8) would not switch to another DMT, but would receive BSC and experience the same rate of disease progression from that point onwards as someone who had the same EDSS and had not received ocrelizumab. To reflect the observations in the MSBase registry, the model allowed the possibility for patients to regress to less severe health states. However, it was assumed that the treatment effect did not directly impact on such regressions. Patients incurred costs and benefits [quality adjusted life-years (QALYs))] as a function of their current EDSS health state 0-9.

In the base-case, utility values for EDSS 0-1 and 8-9 were obtained from a study undertaken by Orme et al. [16]. Briefly, Orme and colleagues undertook a cross-sectional postal survey of patients living with MS in the UK to estimate the disutility of disease progression as well as disutility associated with relapses. Questionnaires were sent to 12,968 patients in the UK MS Trust database, of which 2,048 participants were included in the analysis. Results were presented for patients with different forms of MS. Utility estimates for all other EDSS levels were based on HRQoL information collected using the EQ-5D-3L in the ORATORIO trial. Health-state utility values depended on each health state and thus, were not treatment related. Any disutilities associated with adverse events were obtained from recent technology appraisals and published sources. Carer disutilities by EDSS state were obtained from NICE technology appraisal (TA) TA127. [17]

Costs included in the analysis were those directly related to the NHS. The model included the resource use and treatment costs associated with ocrelizumab. Costs of treatment with ocrelizumab were based on the dose regimen used in the ORATORIO trial using the list price to the NHS (£4,790 per vial). Management costs associated with health state-dependency were obtained from Tyas et al. (2007)[18] and were inflated to current 2017/18 prices. Briefly, Tyas and colleagues undertook a cost analysis and assessed the costs associated with treating patients with multiple sclerosis. Postal surveys with questionnaires were sent to patients in the UK MS database to capture the resource use information associated with the management of MS; then unit costs were assigned to derive the cost per person year from the payers' and societal perspectives. From the 12,698 questionnaires emailed, 2,508 (19.3%) MS participants responded, of which 2,048 (15.8%) were included in the analysis. The analysis included an independent multivariate linear regression of the cost categories using a step-down approach until only statistically significant ($p<0.05$) covariates remained. Cost categories were stratified by direct government-funded costs (direct annual medical/non-medical cost coefficients funded by UK government) and direct out-of-pocket (direct annual medical and non-medical cost coefficients funded out-of-pocket) and indirect costs. Results reported in UK pounds sterling 2005 prices, showed that direct medical costs funded by the government for EDSS levels up to EDSS 4 were not statistically significant from zero, but from EDSS ≥5 reached statistical significance. All non-medical costs funded by the government reached statistical significance for all EDSS levels except zero. Adverse events management costs were obtained from recent technology appraisals. [18] Treatment costs for ocrelizumab were applied until patients discontinued treatment (due to adverse events or progressing to EDSS ≥8), after which it was assumed that patients would incur the costs associated with BSC.

The analysis was undertaken from the NHS and personal social services (PSS) perspective, the outcomes were reported in terms of life-years gained (LYG) and QALYs, with the results were reported in terms of an ICER, expressed as cost per QALY. Both costs and benefits were discounted at 3.5% per annum as

recommended by NICE.[1] Several deterministic one-way sensitivity analyses and scenario analyses were undertaken, as well as probabilistic sensitivity analysis (PSA) based on the outcome cost per QALY only. The company provided results using an agreed patient access scheme (PAS). Under the approved PAS, the company's base-case results showed that the ICER for ocrelizumab versus BSC was estimated at £78,316 per QALY gained in the MRI active population. Sensitivity analysis results showed that the model was sensitive to the treatment effect upon confirmed disability progression at 12 weeks (CDP-12). PSA showed that at a willingness-to-pay (WTP) threshold of £30,000 per QALY gained, ocrelizumab had a zero probability of being cost-effective.

The ERG raised several concerns regarding the values and assumptions in the company's base-case model. Mainly these related to no functionality in the model to allow a waning of the treatment effect, the inclusion of exploratory endpoints (9-HPT and fatigue), which may be used to generate further research and understand the mechanisms rather than be used in the economic model, and the inclusion of utility decrements in the model for upper limb impairment and fatigue. The latter is discussed in detail.

3.2.1. Inclusion of utility decrements

- The company incorporated disutilities to reflect upper limb function using outcomes from the 9-HPT. In the ORATORIO trial, the 9-HPT was included in two outcomes: 20% increase in 9-HPT sustained over 12 weeks and the MSFC. The company chose results for a 20% increase in 9-HPT to reflect upper limb function impairment, hypothesising this corresponds to clinically meaningful upper limb impairment; but made no statement regarding 1) the use of MSFC as a composite outcome that includes the 9-HPT; 2) why MSFC outcomes showed no differences between treatments arms.
- The company incorporated disutilities to reflect fatigue and cognitive impairment as assessed by using the modified fatigue impact scale (MFIS) ≥38. The ERG's understanding is that MFIS denotes how fatigue impacts patients' lives but does not measure cognitive impairment. Cognitive impairment was measured in ORATORIO using the PASAT, and the results showed that there was no statistically significant difference between the treatment arms.
- The company incorporated disutilities related to upper limb and fatigue using the 9-HPT and the MFIS, respectively, which were measured only as part of exploratory analyses in the trial. The ERG was concerned about the selective use of outcomes from exploratory analyses in the base-case of an economic model.

- The ERG considered that there may be potential for double counting of utilities since the EQ-5D adequately captures overall HRQoL for patients with MS. The inspection of the MFIS and EQ-5D questionnaires shows several similarities in the questions. For example, questions pertaining to "self-care" or "usual activities" are captured in the physical subscale of the MFIS as well as EQ-5D. There is also the potential for double counting of utilities using outcomes from the 9-HPT and MFIS. For example, item 4 from the MFIS examines whether patients report "they have been clumsy and uncoordinated". A patient rating "almost always" for this item is also likely to have a poorer score on the 9-HPT. Lastly, some of the MFIS items appear to be linked to progression through the EDSS. As an illustration, a patient responding "almost always" on the MFIS item 13 "my muscles have felt weak" is likely to experience ambulation impairment.
- In addition to utility decrements associated with upper limb, and fatigue and cognitive impairment, the company included carers' disutilities for all EDSS states. Given that the company included utility decrements for caregivers' burden, the ERG considered these additional decrements for upper limb impairment and fatigue to be double counting the impact on QALYs.
- To the ERG's knowledge, utility decrements for upper limb, fatigue and cognitive impairments have not been used in other MS technology appraisals. It was emphasised in the Roche submission that upper limb function is an important outcome for patients with PPMS. The ERG agreed that at high EDSS levels, progression on the EDSS may be poorly responsive to changes, since this scale is heavily dependent on mobility, whilst changes in the manual dexterity performance could be more easily targeted, but it would have been useful to identify the contribution to the observed 9-HPT changes coming from patients with lower EDSS levels and those with higher EDSS scores. Moreover, the ERG noted that this outcome was not incorporated in Roche's submission for ocrelizumab in RRMS.

3.2.2. Regarding the hazard ratio and disutilities derived from 20% increase in the 9-HPT

- A hazard ratio of 0.52 was presented based on the 12-week 9-HPT. The hazard ratio would preferably have been based on the 24-week sustained 20% increase in 9-HPT, which was not reported by the company.
- It appeared that this hazard ratio was derived from patients with EDSS 2 to 6 but was applied to patients with EDSS ≥7. It was unclear whether this hazard ratio generalises to patients in lower (0-1) and higher (≥7) EDSS states.
- There is a lack of transparency about the number of patients randomised to ocrelizumab who experienced a 12-week sustained 20% increase in 9-HPT. Results were only presented for each

EDSS level for the placebo group. For time to 20% increase in 9-HPT, it appears that the hazard ratio was used in the model as a relative risk.

- It was a questionable procedure to incorporate utility decrements into the model when these are based on a 20% increase in 9-HPT.
- The ERG believe that the model should include a feature to allow a waning of the benefit consistent with that using a CDP measure, which is not currently the case.

3.2.3. Regarding the relative risk and disutilities derived for fatigue

- The modified fatigue index score (MFIS) was used to measure fatigue, with a score ≥38 equating to clinically meaningful fatigue. The ERG noted that the baseline mean score for fatigue was 41.6 (SD = 17.2), which suggested that the majority patients were already fatigued upon entering the trial. Figures provided by the company suggested that ocrelizumab had no significant impact on fatigue compared to placebo based on MFIS mean changes.
- The proportion of patients who are likely to experience upper limb deficiency, and fatigue and cognitive impairment at each EDSS level was based solely on the company's clinical expert opinion.

3.3. Appraisal by NICE and final guidance

3.3.1.First appraisal by NICE

The NICE appraisal committee B met in June 2018. The committee considered and discussed the key areas of concern identified by the ERG when they assessed the plausibility of the company's base-case results compared to the ERG's base-case results. The committee agreed with several of the ERG's concerns or comments, in particular the committee concluded that CDP sustained for 24 weeks was preferable to that at measured at 12 weeks and that it was not appropriate to include additional utility decrements for upper limb dysfunction and fatigue. The committee additionally concluded that ocrelizumab's treatment efficacy may likely wane over time, but that the absolute rate of waning was unknown. Considering the committee's preferred assumptions with either the patient access scheme price or the proposed commercial arrangement for the price of ocrelizumab, the committee's appraisal consultation document concluded that ocrelizumab was not likely to be cost-effective.

3.3.2. Second submission by the company, corresponding ERG critique, and appraisal by NICE

Subsequent to release of the consultation document the company submitted new analyses that included fresh evidence regarding the treatment effect size. In their original submission, the company presented post-hoc analyses based on an extended (beyond the period of the double-blinded study) controlled treatment period that added approximately three months of controlled follow-up. The additional period extended from the clinical cut-off date (24 July 2015) to 20 January 2016 or the time when the patient received their first open-label dose of ocrelizumab, whichever came first. During this time, patients were gradually unblinded and switched to an open-label extension (OLE). In the original ERG report, the ERG indicated an interest in results from mature data than observed from the double-blinded period, but cautioned the results might be at risk of performance bias from unblinding.

In the new analyses Roche presented data from the open-label extension study with the data cut-off now extended to Week 336 (5th February 2018).

Although the ERG appreciated the value from this more mature data beyond the OLE study, original concerns regarding the paucity of concurrent randomised control data for comparison, the risk of performance bias applied here to a greater extent since all patients in the ocrelizumab arm were eventually aware of their treatment allocation. There was also the potential for detection bias as outcome assessors may have become aware of the treatment allocation.

The ERG noted that in the original submission the company preferred to use the effectiveness estimates from the double-blind period over those from the extended controlled treatment period. Conversely in the re-submission, the company used estimates from the OLE study in the base-case analysis. The ERG indicated that these OLE study estimates were at higher risk of bias, hence the ERG's preferred base-case used efficacy estimates for un-extended treatment-controlled period (minimum of 120 weeks of double-blinded controlled period).

The company submitted revised base-case results that were partially based on implementing the committee's preferred inputs. However, this revised base-case included inputs for health state utility values and utility decrements for upper limb impairment; assumptions based on the company's preferences. Additionally, the company undertook scenario analyses and assessed the impact of these changes on the deterministic results. The company stated that results were sensitive to the source of treatment efficacy, health state utility values, utility decrements for upper limb impairment, waning of the treatment effect and treatment duration.

A further AC meeting took place in August 2018. In its revised base-case submitted at consultation, and under the agreed patient access scheme, the ICER determined by the company was at £62,766 per QALY

gained for the MRI-active subgroup. However, the ICER rose even higher when based on the committee's preferred assumptions:

- Using efficacy data for CDP-24 from the un-extended controlled period of ORATORIO (rather than including OLE)
- Excluding utility decrements for upper limb impairment
- Excluding utility decrements for fatigue
- Including non-medical direct costs
- Using MRI active subgroup
- Treatment waning – assuming treatment waning starts either from 10 years (as per company's revised base-case) or from 7 years (as per ERG updated base case)
- Including costs, disutilities and treatment effect related relapses
- Using utility values for EDSS from ORATORIO supplemented with utility values obtained from Orme et al.[16]
- Using the updated direct health state costs used by Roche in their revised base-case (submitted at consultation)
- Stopping rule from EDSS 7 (of note, there were a few patients in the trial with EDSS 7, however changing stopping rule from EDSS 7 to 8 had little impact of the ICER)
- Annual discontinuation rates according to company analyses post-ACD
- Including progressive multifocal leukoencephalopathy (PML) as an adverse event, considering its known association with rituximab[19]
- Appropriate to use MSBase registry to inform baseline transitions between EDSS states

The final appraisal document (FAD) published by NICE in September 2018 concluded that ocrelizumab was not recommended, within its marketing authorisation, for treating early PPMS with imaging features characteristic of inflammatory activity in adults.

3.3.3. Third appraisal by NICE and Final guidance

In October 2018, NICE announced that the publication of final guidance for ocrelizumab for PPMS was paused while further discussions took place between the company and NHS England. Based on a revised commercial agreement, the company submitted new analyses to NICE in March 2019. The results of the new analyses were sent to the ERG for verification that the committee's preferred assumptions were

considered as summarized in the September 2018 FAD. A third AC meeting was held in March 2019, after which it was concluded that ocrelizumab, within the commercially confidential arrangement, was cost-effective for treating patients with early PPMS, and in June 2019 final guidance stated that ocrelizumab is recommended, within its marketing authorisation, as an option for treating early PPMS with imaging features characteristic of inflammatory activity in adults contingent upon the company providing ocrelizumab according to the agreed commercial arrangement.[20]

4. Conclusion

Compared with other appraisals conducted on drugs used for MS, the context of ocrelizumab appraised for patients with PPMS was different. One reason for this was the absence of any other drug licensed in the indication, which greatly contrasts with RRMS. Consequently, the unmet need for new therapeutics was viewed as significant, and translated into high patient expectations that approval be granted for ocrelizumab within the NHS. Other particularities were that the main source of clinical effectiveness relied on a post-hoc defined population subgroup of the pivotal trial submitted to the EMA, as well as the incorporation of unusual inputs to value utilities in the cost-effectiveness model.

5. Acknowledgements:

The views expressed in this report are those of the authors and not necessarily those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. Any errors are the responsibility of the authors. This report is supported by OC who is a researcher at the National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre.

6. Author Contributions:

PA, JC, MC, EL, RC, OC, CC, and XA were authors of the ERG report on which that this paper is based. PA and XA produced the first draft of the manuscript. All authors commented on the manuscript and approved the final version. This article has not been externally peer reviewed by PharmacoEconomics.

7. Compliance with Ethical Standards

7.1. Funding

This project was funded by the NIHR HTA programme (project number 16/56/22.). See the HTA programme website (http://www.hta.ac.uk) for further project information: https://www.journalslibrary.nihr.ac.uk/programmes/hta/165622/#/

7.2. Conflicts of Interest

PA, JC, MC, EL, RC, CC, and XA have no conflicts of interest that are directly relevant to the content of this article. OC is a NIHR Research Professor and has received grants from UK MS Society, National MS Society, NIHR-HTA, Rosetrees Trust and served as a consultant for Roche, Merck, Teva, Novartis, Biogen. She is an Associate Editor of Neurology for which she receives an honorarium.

8. References

1. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013. Available from: https://www.nice.org.uk/process/pmg9/. Accessed 20 March 2018.
2. Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. Lancet. 2018 Apr 21;391(10130):1622-36.
3. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014 Jul 15;83(3):278-86.
4. Miller DH, Leary SM. Primary-progressive multiple sclerosis. Lancet Neurol. 2007 Oct;6(10):903–12.
5. National Institute for Health and Care Excellence. Lower price for MS drugs paves the way for positive recommendation from NICE. 2019. Available from: https://www.nice.org.uk/news/article/lower-price-for-ms-drug-paves-the-way-for-positive-recommendation-from-nice. Accessed 25 September 2019.
6. National Clinical Guideline Centre. Multiple Sclerosis: Management of Multiple Sclerosis in Primary and Secondary Care. Clinical guideline 186. 2014. Available from: http://www.nice.org.uk/guidance/cg186/evidence/full-guideline-193254301. Accessed 10 April 2018.
7. European Medicines Agency (EMA). European Public Assessment Report: Ocrevus. 2017. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004043/WC500241126.pdf. Accessed 19 March 2018.
8. National Institute for Health and Care Excellence. Ocrelizumab for treating relapsing–remitting multiple sclerosis: Technology appraisal guidance [TA533]. 2018. Available from: www.nice.org.uk/guidance/ta533. Accessed 29 August 2019.
9. Hawker K, O'Connor P, Freedman MS, Calabresi PA, Antel J, Simon J, et al. Rituximab in patients with primary progressive multiple sclerosis: results of a randomized double-blind placebo-controlled multicenter trial. Ann Neurol. 2009 Oct;66(4):460–71.
10. Montalban X, Hauser SL, Kappos L, Arnold DL, Bar-Or A, Comi G, et al. Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. N Engl J Med. 2017 Jan 19;376(3):209-20.
11. Lublin F, Miller DH, Freedman MS, Cree BAC, Wolinsky JS, Weiner H, et al. Oral fingolimod in primary progressive multiple sclerosis (INFORMS): a phase 3, randomised, double-blind, placebo-controlled trial. Lancet. 2016 Mar 12;387(10023):1075-84.
12. Feys P, Lamers I, Francis G, Benedict R, Phillips G, LaRocca N, et al. The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. Mult Scler. 2017 Apr;23(5):711-20.
13. Institute for Clinical and Economic Review. Disease-Modifying Therapies for Relapsing Remitting and Primary-Progressive Multiple Sclerosis: Effectiveness and Value. 2017. Available from: https://icer-review.org/wp-content/uploads/2016/08/CTAF_MS_Final_Report_030617.pdf. Accessed 19 April 2018.
14. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983 Nov;33(11):1444–52.
15. Butzkueven H, Chapman J, Cristiano E, Grand'Maison F, Hoffmann M, Izquierdo G, et al. MSBase: an international, online registry and platform for collaborative outcomes research in multiple sclerosis. Mult Scler. 2006 Dec;12(6):769-74.

16. Orme M, Kerrigan J, Tyas D, Russell N, Nixon R. The effect of disease, functional status, and relapses on the utility of people with multiple sclerosis in the UK. Value Health. 2007 Jan-Feb;10(1):54–60.
17. National Institute for Health and Care Excellence. Natalizumab for the treatment of adults with highly active relapsing–remitting multiple sclerosis. 2007. Available from: www.nice.org.uk/guidance/ta127. Accessed 22 March 2018.
18. Tyas D, Kerrigan J, Russell N, Nixon R. The distribution of the cost of multiple sclerosis in the UK: how do costs vary by illness severity? Value Health. 2007 Sep-Oct;10(5):386-9.
19. Focosi D, Tuccori M, Maggi F. Progressive multifocal leukoencephalopathy and anti-CD20 monoclonal antibodies: What do we know after 20 years of rituximab. Reviews in medical virology. 2019 Aug 1:e2077.
20. National Institute for Health and Care Excellence. Final appraisal document: Ocrelizumab for treating primary profgressive multiple sclerosis. 2019. Available from: https://www.nice.org.uk/guidance/ta585/documents/final-appraisal-determination-document-2. Accessed 3 January 2020.

Table 1: summary of results from the ORATORIO trial on the main endpoints related to disability progression

Description of the population	**ITT population**		**MRI active patients (T1 Gd-enhancing at screening/baseline or new T2 lesions between screening and baseline)**	
Matching with the label indication at baseline	NO (only partial)		YES (as defined by the Company)	
Type of analysis with regards to the population	Pre-specified / powered		Post-hoc analysis / unpowered	
Arms (number of patients)	Ocrelizumab (n=488)	Placebo (n=244)	Ocrelizumab (n=189)	Placebo (n=104)
Pre-specified primary analysis (clinical cut-off date after a minimum of 120 weeks of double-blind controlled follow-up)				
Patients with 12-week CDP	32.9%	39.3%	32.8%	43.3%
HR for 12-week CDP (95% CI); p-value (log-rank)	0.76 (0.59, 0.98); p=0.0321		0.68 (0.46,0.99); p=0.0448	
Patients with 24-week CDP	29.6%	35.7%	30.7%	38.5%
HR for 24-week CDP (95% CI); p-value (log-rank)	0.75 (0.58, 0.98); p=0.0365		0.71 (0.47,1.06); p=0.0917	
Extended controlled treatment period (post-hoc analysis)				
Patients with 12-week CDP	36.3%	43.4%	NR	NR
HR for 12-week CDP (95% CI); p-value (log-rank)	0.74 (0.58, 0.95) p=0.0151		0.69 (0.47, 1.00); p=NR	
Patients with 24-week CDP	31.6%	40.2%	NR	NR
HR for 24-week CDP (95% CI); p-value (log-rank)	0.70 (0.54, 0.90); p=0.0056		0.68 (0.46,0.99); p=NR	

CDP, confirmed disability progression; CI, confidence interval; HR, hazard ratio; ITT, intention-to-treat; MRI, magnetic resonance imaging; NR, not reported

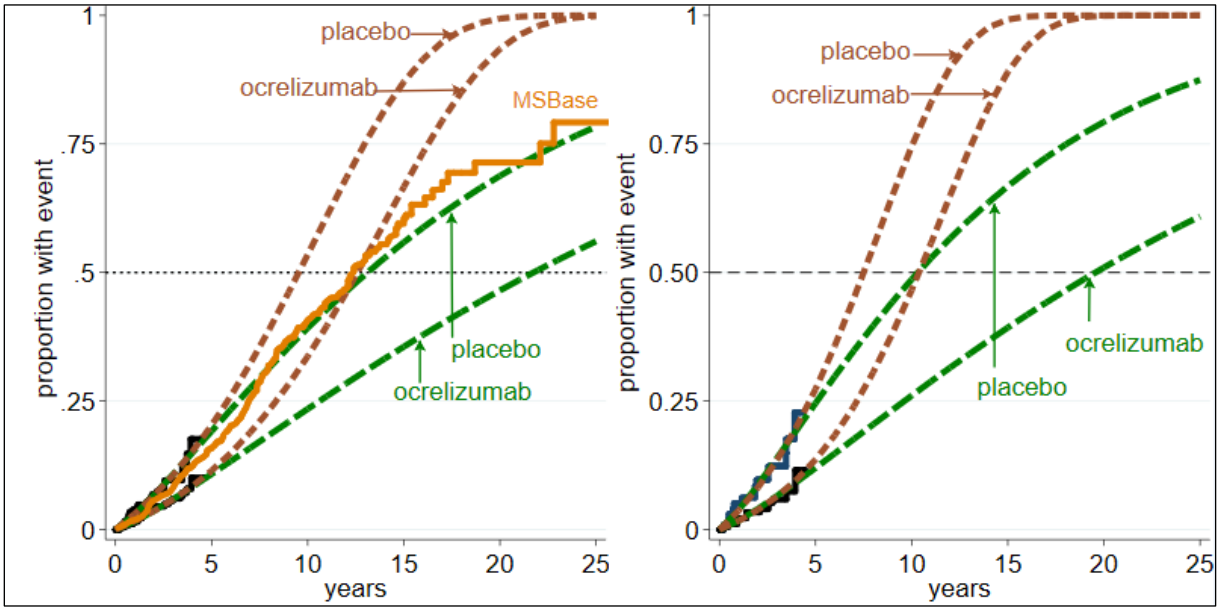

Figure 1: Time to onset of confirmed EDSS ≥ 7 comparison of Weibull (green long dash) and Gompertz (brown short dash) models (left: Intention-to-treat (ITT) population; right: Magnetic resonance imaging (MRI) MRI-active population)