Rituximab in combination with corticosteroids for the treatment of anti-neutrophil cytoplasmic antibody-associated vasculitis

Produced by School of Health and Related Research (ScHARR), The University of Sheffield

Authors Nicholas Latimer, [redacted]

Christopher Carroll, [redacted]

Ruth Wong, [redacted]

Paul Tappenden, [redacted]

Correspondence to Nicholas Latimer, [redacted]

Date completed 30th May 2013

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/37/01.

Declared competing interests of the authors

NL undertook one day of consultancy for Pfizer in 2012, providing advice on the analysis of an unrelated drug in an unrelated disease area. Pfizer manufacture cyclophosphamide and methotrexate (both of which are available on a non-proprietary basis). RL declares that the University of Oxford received an honorarium from Roche Oy for an academic presentation given by RL in 2012.

Acknowledgements

The authors wish to thank Andrea Shippam ([redacted]) for providing administrative support and preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Latimer N, Carroll C, Wong R, Tappenden P, Venning MC, Luqmani R. Rituximab in combination with corticosteroids for the treatment of anti-neutrophil cytoplasmic antibody-associated vasculitis. A Single Technology Appraisal. School of Health and Related Research (ScHARR), The University of Sheffield, 2013.

Contributions of authors

NL led the project, critically appraised the cost-effectiveness section of the manufacturer's submission, scrutinised the manufacturer's model, wrote the model description and conducted the additional ERG analyses. CC critically appraised the clinical effectiveness section of the manufacturer's submission, reviewed additional evidence and contributed to the writing of the report. RW critically appraised and repeated the manufacturer's searches and contributed to the writing of the report. PT assisted with the scrutiny of the manufacturer's model and contributed to the writing of the report. MV and RL provided expert clinical advice, reviewed and contributed to the writing of the report.

Table of Contents

1.	SUMMARY	1
1.1	Summary and critique of the decision problem in the manufacturer's submission	1
1.2	Summary of clinical effectiveness evidence submitted by the manufacturer	2
1.3	Summary of the ERG's critique of clinical effectiveness evidence submitted	2
1.4	Summary of cost effectiveness evidence submitted by the manufacturer	3
1.5	Summary of the ERG's critique of cost effectiveness evidence submitted	4
1.6	ERG commentary on the robustness of evidence submitted by the manufacturer	5
1.7	*Summary of exploratory and sensitivity analyses undertaken by the ERG*	7
1.8	Further issues relating to implementation	7
2.	BACKGROUND	9
2.1	Critique of manufacturer's description of underlying health problem	9
2.2	Critique of manufacturer's overview of current service provision	10
3.	CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM	14
3.1	Population	15
3.2	Intervention	17
3.3	Comparators	18
3.4	Outcomes	20
3.5	Other relevant factors	21
4.	CLINICAL EFFECTIVENESS	22
4.1	Critique of the methods of review(s)	22
4.2	Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)	30
4.3	Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison	38
4.4	Critique of the indirect comparison and/or multiple treatment comparison	38
4.5	Additional work on clinical effectiveness undertaken by the ERG	38
4.6	Conclusions of the clinical effectiveness section	40
5.	COST EFFECTIVENESS	44
5.1	ERG comment on manufacturer's review of cost-effectiveness evidence	44
5.2	Description of the manufacturer's economic model	46
5.3	Critique of the manufacturer's economic analysis	74
5.4	Summary of key issues and uncertainties within the manufacturer's model	101
6.	IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG	103
6.1	Additional work on cost-effectiveness undertaken by the ERG	103
6.2	Discussion of the cost effectiveness section	121

7.	End of life	124
8.	Overall conclusions	125
8.1	Summary of clinical effectiveness issues	125
8.2	Summary of cost effectiveness issues	125
8.3	Implications for research	127
9.	APPENDICES	128
Appendix 1	ERG searches	128
Appendix 2	Critical appraisals completed by the ERG	133
10.	REFERENCES	140

TABLES

Table 1	Decision problem (see MS p.31-32)	14
Table 2	Study characteristics	31
Table 3	Study outcomes	32
Table 4	(reproduced and adapted from MS)	34
Table 5	(reproduced and adapted from MS)	35
Table 6	Number of adverse events in any arm	35
Table 7	Possible RCTs for indirect comparison with CYC as comparator	39
Table 8	Relevant ongoing indirect comparison RCTs	39
Table 9	Initial patient characteristics for patient subgroups and the overall population	54
Table 10	Model parameters and distributions (all subgroups unless indicated)	55
Table 11	One-way sensitivity analyses considered	68
Table 12	Headline cost-effectiveness results presented by the manufacturer	70
Table 13	Structural sensitivity analysis cost-effectiveness results presented by the manufacturer	72
Table 14	Structural sensitivity analysis cost-effectiveness results presented by the manufacturer	74
Table 15	Adherence to the NICE Reference Case	75
Table 16	Correcting mistakes in the manufacturer's economic model	104
Table 17	Tests and costs	106
Table 18	Outpatient appointments	106
Table 19	Altering parameters in the manufacturer's economic model	107
Table 20	ERG preferred analyses cost-effectiveness results – All patients	108
Table 21	ERG preferred analyses cost-effectiveness results – Treatment naïve patients	110
Table 22	ERG preferred analyses cost-effectiveness results – Recurrent disease subgroup (eligible for additional cyclophosphamide treatment)	112

Table 23	ERG preferred analyses cost-effectiveness results – Recurrent disease subgroup (ineligible for additional cyclophosphamide treatment)	114
Table 24	ERG preferred analyses cost-effectiveness results – Cyclophosphamide intolerant subgroup	116
Table 25	ERG additional scenario analysis – All patients: ICERs	120

FIGURES

Figure 1	Rituximab model structure (drawn by the ERG)	47
Figure 2	Probability of remaining in remission (Figure 14 taken from MS)	88
Figure 3	Cost effectiveness acceptability curves – All patients analysis	109
Figure 4	Cost effectiveness acceptability frontier – All patients analysis	109
Figure 5	Cost effectiveness acceptability curves – Treatment naïve subgroup	111
Figure 6	Cost effectiveness acceptability frontier – Treatment naïve subgroup	111
Figure 7	Cost effectiveness acceptability curves – Recurrent disease subgroup (eligible for cyclophosphamide)	113
Figure 8	Cost effectiveness acceptability frontier – Recurrent disease subgroup (eligible for cyclophosphamide)	113
Figure 9	Cost effectiveness acceptability curves – Recurrent disease subgroup (ineligible for cyclophosphamide)	115
Figure 10	Cost effectiveness acceptability frontier – Recurrent disease subgroup (ineligible for cyclophosphamide)	115
Figure 11	Cost effectiveness acceptability curves – Cyclophosphamide intolerant subgroup	117
Figure 12	Cost effectiveness acceptability frontier – Cyclophosphamide intolerant analysis	117

List of abbreviations

AAV	ANCA-Associated Vasculitis
ADCC	Antibody Dependent Cellular Cytotoxicity
AEs	Adverse Events
AIC	Akaike's Information Criterion
ANCAs	Anti-neutrophil Cytoplasmic Antibodies
AUC	Area Under the Curve
BIC	Bayesian Information Criterion
BSA	Body Surface Area
BSC	Best Supportive Care
BSR	British Society for Rheumatology
BVAS	Birmingham Vasculitis Activity Score
BVAS/WG	Birmingham Vasculitis Activity Score for Wegener's Granulomatosis
CEAC	Cost Effectiveness Acceptability Curve
CEAF	Cost Effectiveness Acceptability Frontier
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
CRP	C-reactive Protein
CYC	Cyclophosphamide
DVT	Deep Vein Thrombosis
ERG	Evidence Review Group
ESR	Erythrocyte Sedimentation Rate
EUVAS	European Vasculitis Study Group
EQ-5D	EuroQol-5D
GC	Glucocorticoids
GFR	Glomerular Filtration Rate
GPA	Granulomatosis with Polyangiitis
HRQoL	Health Related Quality of Life
ICER	Incremental Cost Effectiveness Ratio
ITT	Intention to Treat
IV	Intravenous
LFT	Liver Function Test
MMF	Mycophenolate Mofetil
MPA	Microscopic Polyangiitis
MS	Manufacturer's submission
MTC	Mixed Treatment Comparison
MTX	Methotrexate
PD	Prednisone / Prednisolone
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
QALY	Quality-Adjusted Life Year
RCT	Randomised Controlled Trial
RTX	Rituximab
SF-36	Short Form-36
SmPC	Summary of Product Characteristics
SMR	Standardised Mortality Ratio
VBA	Visual Basic for Applications
VDI	Vasculitis Damage Index
VTE	Venous Thromboembolism
WCI	Worst Case Imputation
WG	Wegener's Granulomatosis
WOCF	Worse observation carried forward

1. SUMMARY

1.1 Summary and critique of the decision problem in the manufacturer's submission

The decision problem is appropriate given the anticipated licensed indication for rituximab in granulomatosis with polyangiitis (GPA) and microscopic polyangiitis (MPA). GPA and MPA are two major forms of systemic vasculitis associated with the presence of anti-neutrophil cytoplasm antibodies (ANCAs) which have comparable clinical features and treatment responses, with subtle differences, GPA having higher incidence in the UK. However, the decision problem presented by the manufacturer does not fully match that described in the final NICE scope, due to the anticipated license. The decision problem described by the manufacturer can be summarised as follows. The relevant population includes patients with severe MPA or GPA vasculitis. The intervention, rituximab, is a chimeric murine/human monoclonal antibody that binds specifically to the CD20 antigen expressed on the surface of B cells; it does not bind to hematopoietic stem or CD20-negative precursor cells. Rituximab (RTX) depletes peripheral B cells by several potential mechanisms, including complement-mediated lysis, antibody dependent cellular cytotoxicity (ADCC)-mediated killing, and apoptosis. It is currently licensed for use for several indications. Rituximab received a Positive Opinion for the MPA and GPA indication from the Committee for Medicinal Products for Human Use (CHMP) in March 2013.[1] The relevant RTX dose is 375mg/m^2 once per week for four weeks – this represents the dose to be included within the anticipated license, although a dose of two times 1g on day 1 and day 15 of the treatment cycle is more commonly used in England and Wales. RTX is considered as a treatment for inducing remission in patients with severe MPA or GPA vasculitis. The relevant comparator is cyclophosphamide (CYC) – other potentially relevant comparators (azathioprine - AZA, methotrexate - MTX, and mycophenolate - MMF) were included in the final NICE scope but the manufacturer deems these to be unsuitable for the severe MPA or GPA vasculitis population. The outcomes in the clinical section are appropriate and include all those specified in the NICE scope with the exception of duration of remission.

There is a disconnect between the outcome data reported in the clinical effectiveness and cost-effectiveness sections. The clinical section presents data on proportions of patients who achieve partial and complete remission following induction therapy, and presents data on proportions of patients who go on to experience limited and severe disease relapses. However in the manufacturer's economic analysis, different levels of treatment response and different relapse severities are not modelled. However, the data used to inform the effectiveness parameters in the economic model are reported in the clinical section, where these were available from the relevant clinical trials.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

1.2.1 Clinical effectiveness of rituximab

The manufacturer's submission (MS) identified two RCTs comparing RTX with CYC as induction therapy for adults with severe ANCA-associated vasculitis (AAV) (RAVE and RITUXVAS). There are no head-to-head RCTs which directly compare RTX with other potentially relevant comparators, such as methotrexate (MTX) or mycophenolate mofetil (MMF) for induction of generalised, "severe" AAV. The submission did not include a meta-analysis or narrative synthesis and thus reports the results (equivalence or superiority) as they are reported in the published studies. The two trials appear to be at a low risk of bias, according to assessment by the Cochrane risk of bias tool and the non-inferiority trial extension of the CONSORT statement (for the RAVE trial).[2,3] The evidence suggests that RTX at 4x375mg/m^2 is non-inferior to oral CYC in terms of induction of remission in adults with AAV and *de novo* disease, and superior to oral CYC in terms of remission in adults with generalised, "severe" AAV who have relapsed one time on CYC (RAVE). The evidence also suggests that RTX at 4x375mg/m^2 plus 2-3 IV pulses of CYC is non-inferior to IV pulse CYC in terms of remission in adults with generalised, "severe" AAV and *de novo* disease (RITUXVAS). The evidence only relates to induction of remission with these specific regimens in adult populations with generalised, "severe" AAV and only the 4x375mg/m^2 dose is considered to be relevant, due to the anticipated license.

1.2.2 Safety of rituximab

The MS identified two RCTs comparing RTX with CYC as induction therapy for adults with generalised, "severe" AAV (RAVE and RITUXVAS) and only presented evidence from these two trials with reference to the safety profile of RTX. According to this evidence, both RTX and RTX plus IV pulse CYC appear to have a similar safety profile to oral CYC and IV pulse CYC. Some questions about rates of infection and malignancies in the RTX arms of the two trials have been expressed in the literature.[4,5]

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The submission consists of a systematic review of the effectiveness of RTX for AAV and a summary of relevant adverse events data. The systematic review identified 2 relevant, published RCTs comparing RTX with CYC for generalised, "severe" AAV (RAVE and RITUXVAS). The effectiveness review is inadequately reported in parts. The inclusion and exclusion criteria are not consistent with the NICE scope and final decision problem; the restricted scope of the submission was justified by the manufacturer based on the limitations of the anticipated licence. The searches were inadequate and poorly reported, although this issue was corrected by the manufacturers within their response to the ERG's clarification questions. Methods of study selection and data extraction were not reported by the manufacturer. The choice of critical appraisal checklist and criteria is not described or justified. The results section of the MS consisted simply of a description or overview of each RCT

and its results. There was no meta-analysis or narrative synthesis. The RITUXVAS trial is presented as supporting evidence only and RAVE is considered the only "pivotal" trial. However, clinical advice received by the ERG, the NICE scope and decision problem criteria, suggest that both trials are relevant and that the two trials should be given equal consideration. The submission presents the results of subgroup analyses, but there is no published, dateable evidence that these were pre-specified and a number of the analyses do not relate directly to the outcomes specified in the decision problem. No evidence was identified by the manufacturer comparing RTX with other relevant comparators identified in the NICE scope, such as MMF or MTX. In the absence of head-to-head trials comparing RTX with these other comparators, a search should have been conducted to identify relevant RCTs for an indirect comparison. The ERG identified a number of published (and ongoing) trials evaluating MMF and MTX that were potentially appropriate for indirect comparison.

The submission provides a separate summary of the safety profile of RTX for AAV. This was neither conducted nor reported as a systematic review. However, this section should have either formed part of the efficacy systematic review or should have been a separate, adverse events systematic review. No details were given of the inclusion or exclusion criteria, or the study selection or data extraction processes for this section the adverse events review. No critical appraisal of the safety evidence was conducted beyond the assessments of the RAVE and RITUXVAS trials undertaken for the efficacy review. Results consisted only of a simple description or overview of the RAVE and RITUXVAS trial evidence. There was no attempt to meta-analyse or synthesise this evidence.

1.4 Summary of cost effectiveness evidence submitted by the manufacturer

The manufacturer submitted a health economic model to estimate the incremental costs and health effects of a treatment sequence beginning with RTX compared to a treatment sequence beginning with CYC for the treatment of severe MPA or GPA vasculitis. Cost-effectiveness is presented in terms of the incremental cost per quality-adjusted life year (QALY) gained from the perspective of the NHS and Personal Social Services (PSS) over a lifetime horizon. The manufacturer's base case analysis presented results for "all patients", but subgroup analyses were also presented for "treatment naïve" patients, patients with "recurrent disease" and patients who were "cyclophosphamide intolerant." The manufacturer's model takes the form of a Markov model including health states for "non-remission", "complete remission", "uncontrolled disease" and "dead". The flow of patients through the model depends upon the sequence of treatments received – more lines of therapy are available in the RTX arm of the model, which delays the transition to "uncontrolled disease" and death (although impacts upon death are minimal).

Based upon an amended version of the manufacturer's model submitted as part of the manufacturer's response to the clarification questions, the modelled RTX treatment sequence is expected to produce

more QALYs than the CYC treatment sequence. Similarly, the RTX treatment sequence is expected to be more expensive than the CYC treatment sequence. Based upon the manufacturer's base case analyses, the incremental cost-effectiveness of the RTX treatment sequence compared to the CYC treatment sequence is expected to be around £8,544 per QALY gained for "all patients", £55,175 per QALY gained for "treatment naïve" patients, and £43,003 per QALY gained for patients with "recurrent disease". For "cyclophosphamide intolerant" patients, RTX is expected to represent a dominant treatment strategy – producing additional QALYs and lower costs than a "best supportive care" comparator.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The manufacturer's model has been implemented generally in line with NICE's Reference Case, however there are deviations particularly with respect to the comparators and the outcomes included in the model. The ERG critiqued the model with respect to its structure and the use of evidence to inform the model's parameters; this highlighted only a small number of technical errors, but identified several parameter values that appear inappropriate. The ERG also identified that not all relevant treatment sequences were included in the manufacturer's model, and that the sequences considered in the manufacturer's base case analyses were not appropriate. The ERG believes the following to represent the most important issues and concerns regarding the manufacturer's submitted economic analysis:

- Several realistic treatment sequences were not modelled for the "all patients" analysis or the subgroup analyses.
- Inappropriate costs and (to a lesser extent) utilities were assumed for the "uncontrolled disease" health state (which could be more accurately described in most cases as "grumbling disease").
- An inappropriate assumption was made that all disease flares lead to immediate re-induction therapy – likely leading to a significant over-estimate of the relapse rate and in most circumstances (depending upon the use of maintenance therapy to achieve drug maintained disease control) an unrealistically quick transition to the "uncontrolled disease" health state.
- Assumptions around the resource use costs associated with the "remission" and "non-remission" health states are questionable – the resource use assumed in the "non-remission" state in particular seems to be considerably over-estimated.
- Inappropriate assumptions were made around weight, body surface area (BSA) and wastage. Weight and BSA seem to be underestimated, and wastage is not included in the base case analyses.
- The manufacturer assumed that the glucocorticoid prednisone would be given alongside CYC or RTX, rather than prednisolone. In a UK context, this is inappropriate.

- The manufacturer considerably over-estimated the amount of oral CYC used in a typical treatment course.
- Several important uncertain parameters were not included in the probabilistic sensitivity analysis conducted by the manufacturer.

The most important of these issues relates to the treatment sequences modelled by the manufacturer. In particular the manufacturer assumed that: i) a second course of RTX would be given to patients who initially did not respond to RTX therapy – despite there being no clinical evidence for this; ii) patients achieving remission after RTX treatment would receive no maintenance therapy; iii) treatment sequences were not modelled in which RTX was given as an induction therapy after CYC failure. Clinical advice received by the ERG suggest that each of these assumptions are inappropriate, and failing to consider all relevant sequences is likely to produce misleading conclusions with respect to the cost-effectiveness of RTX.

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer

1.6.1 Strengths

The submission consists of a systematic review of the clinical effectiveness of RTX for generalised, "severe" AAV and a summary of relevant adverse events data. The inclusion criteria are generally in line with the NICE scope and the decision problem; there are a number of variations that were explained with reference to the anticipated licensed dose of RTX for this indication, i.e. induction of remission only in adult populations with generalised, "severe" AAV, and for the licensed 4x375mg/m^2 regimen only. The MS identified the two relevant RCTs comparing RTX with CYC as induction therapy for adults with generalised, "severe" AAV. No relevant head-to-head RCTs were missed. Relevant efficacy and safety data from the published articles of the RAVE and RITUXVAS trials were extracted and presented in full. Both trials were critically appraised using standard RCT risk of bias criteria and were generally correctly appraised as being at a low risk of bias.

The scope of the manufacturer's economic analysis is generally in line with that proposed by NICE. The ERG found very few technical programming errors within the manufacturer's model. The correction of these identified errors had only a minimal impact upon the results of the economic analysis. The manufacturer's model had the inbuilt flexibility that allowed the ERG to investigate further relevant treatment sequences with minimal reprogramming.

1.6.2 Weaknesses and areas of uncertainty

There are uncertainties concerning efficacy and safety in the specified population because the evidence consists only of two short-duration RCTs using different interventions (RTX or RTX+CYC)

and comparators (by dose and administration) in slightly different populations and only for induction of remission. The evidence base is therefore very limited in size and scope.

It is uncertain if RTX alone will demonstrate equal efficacy and safety in other adult populations or children. There are two possible dosing regimens for induction of remission in this population: 4x375mg/m^2 (the licensed dose) and 2x1g (the commonly used dose). The effectiveness evidence submitted only concerns the 4x375mg/m^2 dose for induction therapy for patients with *de novo* disease or CYC-relapsed generalised, "severe" AAV patients. It is uncertain whether the 2x1g dose would demonstrate equivalent or superior efficacy and/or safety for this population. The question whether RTX or RTX+CYC offers the optimum intervention is also uncertain and has been raised in the literature.

The submission focuses on RTX against CYC as a comparator. It is uncertain whether RTX without CYC would demonstrate equal efficacy and safety if it was compared with IV CYC, which might have a better safety profile and which clinical advice suggests is used more often in expert practice in the UK. It is also uncertain how RTX compares with other relevant comparators specified in the scope and decision problem, such as MMF or MTX, which might be used for similar populations (adult patients with generalised, "severe" AAV with and without severe renal impairment respectively). Some evidence is available for an indirect comparison to quantify this uncertainty, however the MS did not contain such an analysis. The long term efficacy and safety of RTX is also uncertain in the specified population beyond 6-18 months.

Finally, the trials appear to be at a low risk of bias, according to assessment by Cochrane and CONSORT criteria for superiority and non-inferiority trials, but there are some outstanding questions regarding the pre-specified nature of some of the analyses presented in the RAVE trial and the reference trial used to establish the non-inferiority criteria. In addition, there is a risk of performance and detection bias in the open-label RITUXVAS trial.

The critical appraisal identified a number of weaknesses within the manufacturer's economic analysis. Key model parameter values (in particular, health state resource use parameters) appear to be highly questionable, and the treatment sequences modelled by the manufacturer do not represent all those that are clinically plausible and valid. The model does not include all relevant health states or comparators, and complete treatment sequences are not modelled. The model is not fully probabilistic, and several important parameters are included only as fixed point estimates. Headline cost-effectiveness results are presented by the manufacturer based upon point estimates of parameters rather than as the expectation of their mean. Given these weaknesses, substantial uncertainty remains

over whether RTX represents a cost-effective addition to clinically relevant treatment sequences and if so, where in the treatment sequence it should be positioned.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG amended the manufacturer's model in order to produce clinically plausible analyses. Identified technical errors were fixed, and seemingly inappropriate parameter values were replaced with more clinically plausible estimates. The most important of these related to the costs associated with the "uncontrolled disease" health state. The ERG amended the manufacturer's model in order to analyse appropriate treatment sequences. In particular, the assumption that some patients received two courses of RTX treatment was removed, an assumption was added that patients induced into remission through RTX treatment went on to receive AZA maintenance therapy, and the inclusion of RTX at different points in the treatment pathway was considered. This allowed the ERG to address two important questions: i) does the inclusion of RTX in the treatment sequence increase health benefits compared to the current treatment sequence? and ii) if so, where is the most cost-effective place in the pathway to position RTX?

These analyses suggest that adding RTX to the standard treatment sequence is associated with an incremental cost effectiveness ratio (ICER) of approximately £10,699 to £12,851 per QALY gained – provided RTX is only used after CYC treatment has been exhausted. Moving RTX forwards in the treatment sequence (as first- or second-line treatment in "treatment naïve" patients who are able to receive two courses of CYC, or as first-line treatment in "recurrent disease" patients who are only able to receive one course of CYC) is associated with much higher ICERs – ranging from £50,842 to £317,038 per QALY gained for different patient groups and treatment sequences. In patients who are intolerant to CYC, or unable to take further CYC due to a high lifetime cumulative dose, the ICER associated with RTX compared to "best supportive care" is expected to be approximately £10,699 to £11,277 per QALY gained, although these analyses are limited and may be biased in favour of RTX due to the exclusion of potentially relevant comparator treatments such as MMF.

Given the limitations associated with the manufacturer's model, considerable uncertainty remains around the cost-effectiveness of RTX.

1.8 Further issues relating to implementation

A number of issues relating to implementation of RTX treatment were raised by the clinical advisors to the ERG:

- **Rituximab as maintenance treatment, and for treating disease flares.** RTX is currently used off-label for the treatment of patients with vasculitis – for induction, maintenance, and

the treatment of disease flares. Ongoing trials are investigating the use of RTX in these indications. This appraisal only addresses the use of RTX as a remission induction therapy. The clinical advisors to the ERG highlighted that there may be practical difficulties associated with inducing remission with RTX and then not providing further RTX treatment for disease maintenance and the treatment of disease flares.

- **Target population.** The decision problem addressed in this appraisal focusses upon patients with severe GPA or MPA vasculitis. However, clinical advisors to the ERG noted that RTX is used off-label for patients with lower severity and for patients with other types of vasculitis. In addition, the clinical advisors to the ERG were keen to emphasise that RTX is not proven in patients with very severe, life-threatening forms of GPA or MPA vasculitis.

- **Co-prescribing with cyclophosphamide.** One of the key clinical trials of RTX in severe GPA or MPA vasculitis, RITUXVAS, involved combination treatment with CYC. It is anticipated that this will not be included on the RTX license and this combination has not been considered in the economic evaluation presented in the MS. However, this may need to be considered in future – the clinical advisors to the ERG suggest that this may be particularly important for patients with very severe, life-threatening vasculitis. Co-prescribing with CYC and RTX may also prevent future damage, by providing more rapid early control of generalised and severe disease.

2. BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer's description of the extent of the underlying health problem and the risks associated with AAV is relatively brief. It is noted that GPA (also known as Wegener's granulomatosis (WG) and MPA are two major forms of systemic vasculitis associated with the presence of anti-neutrophil cytoplasm antibodies (ANCAs) which have comparable clinical features and treatment responses. It is stated that collectively these are referred to as AAV (ANCA-associated vasculitis). Other types of ANCA-associated vasculitis are not mentioned (for example, eosinophilic granulomatosis with polyangitis (EGPA), though this may be reasonable since the rituximab (RTX) license is likely to only include GPA and MPA populations and these share many clinical features.[6] It is estimated that GPA accounts for approximately half of AAV in Northern Europe, with one third accounted for by MPA and the remainder being EGPA.[6] However a brief description of the relative population sizes of different types of ANCA-associated vasculitis, combined with an explanation of why RTX is only to be indicated for GPA and MPA, may have been useful.

In the MS, the manufacturer states that they estimate the total MPA and GPA population to be 13,000 patients annually, based upon prevalence data from a 2012 epidemiological study (see MS p.20).[7,8] It is further estimated that the number of patients that would be eligible for RTX would be 1,660 annually based upon the proposed license being restricted to severe GPA or MPA and induction therapy only. However, the specific figures (prevalence, incidence, severity breakdown) used to estimate these figures are not stated. Watts *et al.* (2012) estimate that the combined average annual incidence of GPA and MPA in Norfolk, UK between 1988 and 2010 was 17.2 per million people, and estimate that prevalence was 209 per million people at the end of 2008.[7] Based upon these figures and a population of approximately 56 million for England and Wales, the incidence of GPA and MPA combined would appear to be approximately 963 patients per year, with a prevalence of approximately 11,700 patients. It is therefore slightly unclear how the manufacturer arrived at their patient population estimate, or how they transformed this into a relevant patient population for RTX. Market research indicating that approximately 5,000 patients per annum are treated with induction therapy is quoted,(see MS[8] p.20) and the manufacturer states that their evidence suggests that one third of these are defined as having severe disease, but no further details are provided.

The manufacturer's description of the risks associated with AAV is adequate. It is stated that early studies showed high 1- and 2-year mortality rates, but that these have been reduced considerably by the introduction of treatment with cyclophosphamide (CYC). A 5-year survival rate of 82% is quoted by the manufacturer, although it is noted that considerable morbidity is associated both with the disease and its treatment – with 20% of survivors having significant renal disease (see MS p.21).[8,9]

It is noted that disease onset usually occurs at 65-74 years, although it can occur at any age.[8,10] Increasing age and renal involvement at diagnosis are indicative of a poor prognosis. After achieving disease remission the likelihood of relapse varies according to disease-type but is highest in GPA; up to 50% of patients will relapse within 5 years, even with maintenance immunosuppression. Each relapse carries a risk of subsequent critical organ damage.[9,11]

The MS states that therapies used to treat AAV are associated with substantial toxicities that frequently result in severe and permanent patient morbidity and mortality.[8,12] Of particular importance, treating AAV with cyclophosphamide (CYC) can lead to opportunistic infections, bone-marrow suppression, hemorrhagic cystitis, infertility, and cancer, in particular hematopoietic and bladder malignancies (see MS p.21).[8,13,14] It is stated that cumulative exposure to CYC is a significant risk factor for adverse events and mortality (see MS p.22).[8] Due to these risks, the MS states that there are specific situations where the avoidance of CYC treatment is desirable:

- Females who have not yet completed their family and who are at risk of CYC-induced premature menopause.
- Where there has been a previous uroepithelial malignancy.
- Where there is intolerance of CYC due to side-effects or cytopenia.
- Where there is a high risk of infection.

The MS also states that there are substantial morbidities associated with a repeated and prolonged course of glucocorticoids. It is stated that infections are a well-known complication of glucocorticoids, especially in the treatment of vasculitis (see MS p.22),[8,15] and that other known complications of steroid therapy include new-onset diabetes, osteoporosis, avascular necrosis, peptic ulcers and cataracts.[8,16]

2.2 Critique of manufacturer's overview of current service provision

The manufacturer's description of current service provision is adequate although some discussion around specific points is required. The ERG and our clinical advisors suggest that the full range of treatment options for patients are not described within the MS.

As the manufacturer states, there is no NICE guidance on AAV treatment. Instead, the MS focusses upon guidelines for the management of AAV produced by the British Society of Rheumatology (BSR) and the European Vasculitis Study Group (EUVAS).[9,17]

The MS states that management of AAV involves three phases; remission induction, remission maintenance and treatment of relapse. Treatment is tailored according to severity and extent of

disease, and the level of immunosuppression reflects the severity of the vasculitis. CYC is the standard remission induction agent, given either daily orally or through IV pulses and is usually given for 3-6 months, adjusted for age, body weight and renal function. CYC is usually given with oral or IV corticosteroids such as prednisolone or methylprednisolone. For aggressive disease, plasmapheresis can be used in addition to standard therapy.

The MS states that to maintain remission, CYC therapy may be continued for up to 6 months, along with tapered doses of prednisolone. Azathioprine (AZA) or methotrexate (MTX) may be substituted for CYC after successful remission of disease. Mycophenolate Mofetil (MMF) or leflunomide may be used as alternatives if patients are intolerant to AZA or MTX, or due to lack of efficacy.

The MS states that patients may continue maintenance therapy/immunosuppression for up to 5 years. The MS states that relapses are classified as minor or major, according to the absence or presence of threatened vital organ function. Minor relapse is treated with an increase in prednisolone dose whereas major relapse is treated with CYC as in remission induction and an increase in prednisolone; intravenous methylprednisolone or plasma exchange may be considered.[8]

The MS states that disease that is refractory to full dose CYC and prednisolone is rare. More commonly, optimal doses are not tolerated or high cumulative exposure to CYC and prednisolone lead to alternative agents being considered. The MS states that for relapses on AZA or MTX a switch to MMF or leflunomide may be considered.

While the ERG and our clinical advisors broadly agree with the treatment pathway outlined by the manufacturer, some clarification is required. Firstly, although the EUVAS guidelines recommend oral or IV CYC, it is stated that IV treatment may result in higher remission rates and lower risk of side effects, although it may also be associated with a higher relapse rate.[17] IV treatment leads to a lower cumulative dose of CYC per treatment cycle compared to oral treatment. The oral CYC dose recommended by the BSR,[9] 2mg/kg/day, would lead to a total dose of approximately 25,568mg over a 6 month period (for a patient weighing 70kg). The IV CYC dose recommended by the BSR, 15mg/kg every two weeks for three pulses, followed by the same dose administered at 3-week intervals, would lead to a total dose of approximately 10,500mg over a 6 month period (for a patient weighing 70kg). Recognising that high cumulative CYC doses are indicative of increased risks of adverse events, the ERG and our clinical advisors believe that the difference between IV and oral CYC as treatment options is not highlighted adequately by the manufacturer. In particular, the administration of IV CYC rather than oral CYC may allow a greater number of courses of CYC to be given.

In addition, the ERG and our clinical advisors believe that the manufacturer has not adequately considered all treatments that may be used to induce remission in patients with AAV. Although CYC represents the clear treatment of choice (other than RTX) in patients with severe AAV, other options do exist and these are not adequately considered within the MS. For a patient who already has a high cumulative exposure to CYC, clinicians may offer alternatives such as MTX, MMF, leflunomide, higher dose AZA, abatacept or infliximab, or combinations of these. Several of these may not be licensed, but our expert advisors suggest that they are used to control disease in patients who have already received CYC (or are intolerant of it or do not wish to receive it), and play a part in many treatment sequences. Importantly, the particular treatment chosen will depend upon individual patient characteristics. For patients with rapidly progressive severe renal disease, plasma exchange represents a recommended treatment option.[9,17] When considering the AAV disease and treatment pathway as a whole, it is important to recognise these options.

With respect to remission-maintenance, it is notable that EUVAS recommend treatment with a combination of low-dose glucocorticoid and either AZA, leflunomide or MTX (the BSR guidance mentions only AZA and MTX) – and it is recommended that this treatment is continued for at least 18 months.[17] This is of particular importance considering the manufacturer's decision to model a treatment pathway in which patients who achieve remission after treatment with RTX receive no maintenance therapy (see Section 5.2.2). The ERG and our clinical advisors believe that this does not reflect a realistic or current treatment pathway.

With regard to the treatment of relapses, the ERG and our clinical advisors agree with the manufacturer that these are treated in very different ways depending upon their severity and location. The ERG believes that this is important to bear in mind when considering the manufacturer's economic model, which does not differentiate between minor and major relapses (see Section 5.2.2).

The ERGs clinical advisors suggest that the description of the disease pathway given by the manufacturer is misleading. For instance, it is stated that *"Although many patients achieve remission with expert care, they usually experience disease flares when therapy is tapered or discontinued"* (see MS p.25).[8] Our clinical advisors suggest that most patients achieve disease remission and that in fact disease relapse is not inevitable given appropriate maintenance treatment. The remission-maintenance phase of the treatment pathway has at least two components – a component whereby disease control is maintained using drug treatment, first at a standard dose and then at a tapering dose. This is followed by a second component whereby drug-free disease control is maintained. The first and second components may be of differing lengths (sometimes drug treatment is given for more than 5 years), but appropriate maintenance lessens the frequency of flares and also reduces their severity such that

relatively moderate treatment adjustment (rather than re-induction treatment) may be sufficient to regain disease control.

Finally, the ERG acknowledges that RTX is already being used to treat AAV in the NHS, as stated by the manufacturer. In particular, the NHS Commissioning Board have stated that they intend to commission RTX for AAV, for patients with relapsing disease, with primary treatment failure or with adverse reactions or contraindications to CYC.[8,18] As the MS states, the NHS Commissioning Board describe four different situations where RTX would be routinely funded for AAV:[18]

- Initial remission induction agent in newly diagnosed patients where avoiding CYC is desirable
- As a remission induction agent when CYC has not been effective
- As a remission induction agent at time of first relapse
- As a remission maintenance agent.

The ERG also acknowledges the manufacturer's statement that a dosing schedule of 1,000mg of RTX on Day 1 and Day 15 of the treatment cycle is commonly used for the treatment of AAV in the UK. While the ERG's clinical advisors suggest that the use of RTX for treating AAV is not yet widespread in the UK (with the exception of a small number of centres), it is accepted that the 2 times 1,000mg dose is more often used than the four times 375mg/m^2 that forms the basis of the anticipated license indication.

3. CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

A summary of the decision problem defined in the NICE scope[19] and addressed in the MS is shown in Table 1. Where the decision problem addressed is different from that in the NICE scope the rationale for the difference is as given by the manufacturer in the MS (see MS p.31-32).[8]

Table 1: Decision problem (see MS p.31-32)

	Final scope issued by NICE	**Decision problem addressed in the submission**	**Rationale if different from the scope**
Population	People with anti-neutrophil cytoplasmic antibody associated vasculitis	People with severe ANCA-associated vasculitis (GPA or MPA)	In line with the proposed license for rituximab
Intervention	Rituximab in combination with corticosteroids	Rituximab in combination with corticosteroids for induction treatment only.	In line with the proposed license for rituximab
Comparators	Treatment strategies without rituximab, including cyclophosphamide, azathioprine, methotrexate, and mycophenolate (in combination with corticosteroids)	Cyclophosphamide (in combination with corticosteroids) for induction of remission. Patients intolerant to cyclophosphamide may receive alternative induction treatments including azathioprine, methotrexate, and mycophenolate (in combination with corticosteroids)	Clinical opinion has informed us that with the exception of patients intolerant to cyclophosphamide, cyclophosphamide is the standard of care in the induction of remission, with the use of other agents usually reserved for less severe forms of AAV. The RAVE study compared rituximab with cyclophosphamide. We have no evidence of rituximab directly compared to other induction agents in the target population.
Outcomes	The outcome measures to be considered include: • mortality • remission rate and duration of remission • number and severity of relapses • change in renal function • cumulative dose of immunosuppressants • adverse effects of treatment • health-related quality of life.	The following outcome measures were included based on the evidence available: • mortality • remission rate and duration of remission • relapse rates • cumulative dose of immunosuppressants • adverse effects of treatment • health-related quality of life	Not all the outcomes outlined in the final scope could be included due to limitations in the evidence base. The RAVE clinical study has provided information related to relapse rates, however, we have not been able to model the severity of individual relapses due to limitations in the data.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost	Same as the scope	-

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
	effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.		
Subgroups to be considered	People for whom cyclophosphamide is contraindicated	The following subgroups were explored: • Patients with *de novo* disease • Patients with prior exposure to cyclophosphamide • Patients for whom cyclophosphamide is contraindicated	The RAVE study included patients with either newly diagnosed or relapsing disease. The study reported significantly different outcomes based on prior exposure to cyclophosphamide. Due to the limitations in the available evidence we have undertaken scenario analyses for patients who are considered intolerant to cyclophosphamide and are receiving alternative induction agents (azathioprine, methotrexate, and mycophenolate).
Special considerations, including issues related to equity or equality		Cyclophosphamide is known to reduce the fertility of both males and females. Rituximab offers an effective alternative to patients looking at preserving their fertility without compromising on the effectiveness of their treatment.	

3.1 Population

The patient population addressed by the manufacturer's statement of the decision problem differs from that described in the final NICE scope.[19] Whereas the NICE scope intended for the appraisal to address all patients with ANCA-associated vasculitis, the MS only considers patients with what the manufacturer defines as "severe" MPA or GPA vasculitis. The rationale for this is that the anticipated RTX license will only cover this more limited population. In addition, the two pivotal randomised controlled trials (RCTs) upon which the appraisal rests, RAVE and RITUXVAS, only included patients with GPA and MPA AAV. However, the ERGs clinical advisors suggest that other types of ANCA-associated vasculitis may also receive RTX (and in fact do so in clinical practice in England and Wales). It may also be used in less severe AAV in specific patients – for example, women who

wish to become pregnant. However, given the anticipated licensed indication, it seems appropriate that the MS focuses upon "severe" GPA and MPA vasculitis.

The fact that the manufacturer focuses upon a population with "severe" GPA or MPA vasculitis creates further problems, since the definition of severity is not clear-cut, and the population included in the RAVE and RITUXVAS trials differed. In Section 8 of the MS, it is stated that the estimated numbers of patients that would be eligible for RTX treatment are based upon the disease severity classification used in RAVE – the manufacturer assumes that one third of patients have severe disease (see MS p.181).[8] In RAVE, only patients with "severe" disease were recruited – these patients were defined as those with one or more of the major BVAS/WG items, or disease severe enough to require treatment with CYC. However, it is notable that the RAVE trial excluded patients with severe disease who required mechanical ventilation because of alveolar haemorrhage, and patients with a serum creatinine level of greater than 4.0mg/dL attributed to underlying AAV disease – thus the clinical evidence submitted is not relevant for *all* patients with severe disease. The ERG's clinical advisors have suggested that the exclusion criteria applied in RAVE mean that patients with severe renal disease were excluded, as well as patients with other life-threatening forms of the disease (such as life-threatening lung haemorrhage and severe cerebral vasculitis).

The EUVAS and BSR guidelines categorise severe ANCA-associated vasculitis as disease including renal or other vital organ failure, with serum creatinine of greater than 500μ/mol/litre (5.6mg/dL). Hence it seems that these patients would actually have been excluded from the RAVE trial.[9,17] Under these guidelines, the RAVE definition of severe disease appears closer to what is classified as "generalised" disease – where vital organ function is threatened and serum creatinine is less than 500μ/mol/litre (5.6mg/dL). In their response to clarification questions, the manufacturer stated that the definition of severe renal disease is problematic, and that some patients included in RAVE, and particularly in RITUXVAS, may have been regarded as having severe renal disease.[20] However, the ERG's clinical advisors suggest that since patients with serum creatinine of greater than 500μ/mol/litre were excluded from RAVE conclusions cannot be made upon the effectiveness of RTX in this group. Also, RITUXVAS included few of these patients, and administered RTX in combination with CYC – hence this study cannot be used to show that RTX monotherapy is effective in patients with severe renal disease. In this report, the ERG refer to the population included in the RAVE trial as generalised, "severe", but recognise that according to other definitions this population may be classed as having "generalised" disease, and that patients with the most severe disease are excluded.

In their response to clarification questions the manufacturer also stated the following (see manufacturer's Clarification Response,[20] p.27):

"Patients with the most severe forms of pulmonary haemorrhage were excluded from both RAVE and RITUXVAS because mortality is high and physicians are generally reluctant to recruit critically ill patients to this type of trial. However, time to remission for patients with less severe lung haemorrhage was a secondary end-point in RAVE and there was no difference between CYC and RTX groups. Expert opinion believes that there would be no reason to think responses in severe pulmonary haemorrhage would be any different.

Expert practice, and their advice to others, is to use rituximab for the most severe forms of pulmonary haemorrhage because the major cause of death is infection in the second or third intensive care week and they regard the infection risks of RTX and steroids as being easier to manage than those with CYC and steroids."

The clinical advisors to the ERG disagree with this statement. Whilst opinion may be split with respect to the most appropriate treatment for patients with severe pulmonary haemorrhage, the suggestion that this is a majority view appears to be misleading.

There is an additional question as to whether the population included in the RAVE trial (which forms the basis of the manufacturer's economic evaluation) represents a population relevant for the UK population. The manufacturer states that the RAVE study recruited mainly in the United States and the Netherlands, and that the average body surface area (BSA) of included patients was 1.998m^2, with an average weight of 87kg (although other evidence submitted by the manufacturer suggested that the average weight was actually 85.1kg).[21] The manufacturer states that this weight and BSA is not reflective of UK patients, and instead an average BSA of 1.79m^2 and an average weight of 67kg was assumed based upon estimates for UK cancer patients (see MS p.138).[8] This is important due to the weight and BSA-related doses of several of the treatments included in the manufacturer's economic model.

Finally, in the MS it is stated that *"only RAVE… reflects the license and scope of this appraisal"* (see MS p.91).[8] The ERG disagree with this statement, since the RAVE trial only considers a subset of the final NICE scope's population, interventions and outcomes.

3.2 Intervention

Rituximab is a chimeric murine/human monoclonal antibody that binds specifically to the CD20 antigen expressed on the surface of B cells; it does not bind to hematopoietic stem or CD20-negative precursor cells. RTX depletes peripheral B cells by several potential mechanisms, including complement-mediated lysis, antibody dependent cellular cytotoxicity (ADCC)-mediated killing, and apoptosis. It is currently licensed for use for several indications. It received a Positive Opinion for the

MPA and GPA indication from the Committee for Medicinal Products for Human Use (CHMP) in March 2013.[1]

The intervention described in the MS differs slightly from that described in the final NICE scope due to its anticipated license indication. The final scope did not specify whether the intervention would be considered for induction, maintenance, or relapse therapy. However the manufacturer anticipates that the intervention will be licensed only for induction therapy and so only this indication is considered in the MS. This is despite the fact that the RAVE trial included the treatment of flares with RTX, and other clinical trials have studied maintenance treatment with RTX. This causes some problems for the decision problem faced in the appraisal, because the ERG's clinical advisors state that it is very unlikely that patients who achieve remission through treatment with RTX would go on to receive no maintenance therapy. It is more likely that these patients would receive maintenance therapy with RTX or another treatment (such as AZA) – yet in the MS the manufacturer models a treatment pathway in which no maintenance therapy is received by patients induced into remission by RTX (see Section 5.2.2). Our clinical advisors suggest that patients treated successfully with RTX will regain their B-cells over time and are therefore likely to relapse, suggesting that without maintenance therapy disease may be badly controlled. Given that the manufacturer's economic model is based primarily on the RAVE trial which included no maintenance treatment in the RTX arm, and that the anticipated license will not include RTX as a maintenance therapy, this represents an important issue. Given the anticipated license, and advice received by the ERG from our clinical advisors, it would appear appropriate to assume that patients who achieve remission on RTX then receive maintenance therapy with AZA – representing the maintenance treatment received only in the CYC arm of the RAVE trial.

It is also important to note that in the manufacturer's economic model it is assumed that patients who initially do not respond to RTX treatment are immediately given a second course of RTX, offering a further chance of achieving remission. Such retreatment has not been studied in any of the evidence submitted by the manufacturer, but it appears not to be ruled out by the anticipated license. The manufacturer states that they modelled this sequence based upon expert opinion. However advice received from the ERG's clinical advisors is contrary to this – suggesting that there is no evidence in favour of this and that it would be more likely that an alternative therapy would be tried.

3.3 Comparators

The comparators defined in the final NICE scope were *"Treatment strategies without rituximab, including cyclophosphamide, azathioprine, methotrexate, and mycophenolate (in combination with corticosteroids)"*,[19] while the comparator included in the manufacturer's submission is *"Cyclophosphamide (in combination with corticosteroids) for induction of remission. Patients intolerant to cyclophosphamide may receive alternative induction treatments including azathioprine,*

methotrexate, and mycophenolate (in combination with corticosteroids)".[19] In the manufacturer's economic evaluation, the base case analyses compare a treatment strategy that includes two courses of RTX followed by one course of CYC to a treatment strategy that includes two courses of CYC. Other treatments are not included in the treatment strategies for any analyses. Even for the subgroup analysis of patients who are intolerant to CYC, the administration of other treatments was not modelled – it is instead assumed that patients enter an "uncontrolled disease" health state. This appears to be because no direct head-to-head evidence is available comparing RTX to treatments other than CYC. No indirect comparisons against other treatments were undertaken.

While the ERG agrees that CYC represents the most relevant comparator for treatment to induce remission, in a disease such as vasculitis in which multiple relapses and remissions are possible, it is important to accurately model realistic treatment sequences in order to accurately reflect the lifetime impact of adding new lines of treatment, and because treatments may be positioned at different places within the treatment pathway. It is relevant to note that early results of the MYCYC trial (investigating MMF compared to CYC for remission induction in patients with severe or moderate ANCA-associated vasculitis) recently published in abstract form, suggest that MMF provides similar remission rates compared to CYC, although non-inferiority has not yet been proven.[22] Expert clinical advice received by the ERG suggests that the absence of other potentially useful induction treatments in the treatment pathway included in the manufacturer's model as well as the absence of maintenance therapy in the RTX arm, is likely to predispose patients to transit into an "uncontrolled disease" health state much more quickly than is realistic. In fact, our advisors suggest that the "uncontrolled disease" health state may in itself be unrealistic – this is discussed in Section 5.3. However, considering that subsequent induction treatment with interventions such as MTX, MMF or AZA is excluded from both arms of the manufacturer's model the impact may be relatively minor. The ERG suggests that what is likely to be most important is a consideration of whether RTX represents an additional treatment in the treatment pathway, or whether it replaces another treatment. The manufacturer's model implicitly assumes that RTX offers an additional line of therapy, since two courses of RTX may be received followed by one course of CYC – 3 lines of treatment in total, compared to the 2 courses of CYC modelled in the control arm.

Advice received by the ERG from our clinical advisors suggests that RTX is likely to represent an additional treatment option, but that this may be due to patients receiving the same number of courses of CYC over their lifetime, but with the addition of one course of RTX (more courses of RTX may be given over time if initial treatment is successful, but this is not expected to be included in the anticipated license). However, advice received by the ERG also suggests that a key treatment sequence not explicitly modelled by the manufacturer involves initial treatment with CYC, followed by RTX treatment to induce subsequent remission. It appears that this would not be excluded by the

anticipated license, and also reflects the treatment pathway suggested by the NHS Commissioning Board.[18]

Evidence on the use of RTX and CYC (and other potentially relevant comparators) is limited when considering the effectiveness of second, third or subsequent courses of treatment. Some data are available from the RAVE trial with regard to effectiveness in newly diagnosed patients compared to effectiveness in patients with recurrent disease. These two groups of patients each made up approximately half of the RAVE trial population (approximately 50 patients in each randomised treatment group, for each population group). Hence these data are uncertain. However, given the importance of treatment sequences in this appraisal, the ERG believes that these data are important. It is noteworthy that no data are presented by the manufacturer for patients who are intolerant to CYC. Hence the subgroup cost-effectiveness analysis presented for this group is questionable. There may be no *a priori* reason to expect the effectiveness of RTX to differ in patients intolerant to CYC, but in an economic evaluation the comparator should reflect an alternative to CYC such as MTX, MMF or leflunomide rather than no treatment.

3.4 Outcomes

3.4.1 Outcomes in clinical effectiveness section

The outcomes presented in the clinical section of the MS are appropriate and match those included in the NICE scope, with the exception of duration of remission.

3.4.2 Relevance of outcome data to cost-effectiveness model

The outcome data used in the manufacturer's economic model are drawn almost exclusively from 6-month data from the RAVE trial. Data from RITUXVAS are not used (presumably due to the RTX intervention being a combination therapy, which is not considered in the appraisal) and longer term data from the RAVE trial are used only to estimate relapse rates. Not all the outcomes listed in the final NICE scope are incorporated into the economic model. Mortality, remission rates, relapse rates, cumulative dose of immunosuppressants, adverse effects of treatment and health-related quality of life (HRQoL) are all included but different severities of relapse and changes in renal function were not modelled. In addition, mortality rates were not derived from the RAVE trial due to the low number of deaths observed; instead an alternative cohort study data source was used to inform these parameters. Remission rates for first and second courses of CYC treatment were derived from the RAVE trial, however assumptions had to be made regarding the effectiveness of a second course of RTX. Relapse rates were available from the RAVE trial and could be estimated for treatment naïve patients and patients with recurrent disease – these differential rates were tested in the manufacturer's sensitivity analyses and rates based upon data from all patients included in RAVE were applied in the base case analyses (for first and second relapses). Although the cumulative dose of immunosuppressants such as

prednisolone and methylprednisolone observed in the RAVE trial were included in the model, these were assumed to be equal in the two treatment arms.

Advice received from the ERG's clinical advisors suggests that the key weaknesses in the manufacturer's economic evaluation with respect to the inclusion of relevant outcome data involved the failure to model relapses of different severity – since treatment options and the disease pathway depend critically upon whether a minor or major relapse is experienced. Data on relapse severity were available from the RAVE trial and are presented in the clinical section of the MS (for example, see Table 15 on p.68 of the MS).[8] The clinical advisors to the ERG also suggest that the failure to model different levels of treatment response may be important. The manufacturer's model defines four health states – "complete remission", "non-remission", "uncontrolled disease" and "death". Complete remission is defined by a BVAS/WG score of 0 and a successful taper of glucocorticoid therapy at 6-months after randomisation, as defined in the RAVE trial. However, the ERG's clinical experts suggest that a proportion of patients are likely to achieve a BVAS/WG score of 0 but may still be receiving glucocorticoid treatment at 6-months. Data on such patients are presented in the clinical section of the MS – 80.8% of RTX patients and 66.3% of CYC patients achieved a BVAS/WG score of zero, while only 63.6% and 53.1% respectively achieved a BVAS/WG score of 0 *and* a successful glucocorticoid taper at 6-months. Patients who achieve remission without a complete glucocorticoid taper are likely to have a different quality of life and resource use profiles compared to patients who do not achieve a BVAS/WG score of 0, but these patients are classified as non-responders in the economic model and go on to receive a second course of treatment. The difference in the proportion of patients who achieved complete remission and remission without a completed glucocorticoid taper was similar in the RTX and CYC groups in the RAVE trial, and therefore the relative impact of not including a health state for these "non-complete responders" in the economic model may be minor – however the knock-on impacts on subsequent treatments, costs and effectiveness in the economic model is unknown.

3.5 Other relevant factors

The manufacturer's submission does not include discussion of relevant equity considerations.

4. CLINICAL EFFECTIVENESS

This chapter provides an overview and critical appraisal of the clinical effectiveness review submitted by the manufacturer. It presents details of the principal trial evidence and offers a critique of the review in relation to the decision problem and the conduct of each stage of the manufacturer's submitted review, i.e. the searches, the processes of study selection, data extraction, quality assessment (including tools used), and method of synthesis (see Sections 4.1 – 4.1.5). This is followed by the ERG's own critique and commentary on the included trial evidence and an examination of any potentially relevant evidence not covered in the manufacturer's submission. Finally, this section describes the possible conclusions to be drawn from the clinical effectiveness review, and the uncertainties and limitations affecting the evidence.

4.1 Critique of the methods of review(s)

The submission consists of a systematic review of the effectiveness of RTX for AAV and a summary of relevant adverse events data. The systematic review identified 2 relevant, published RCTs comparing RTX with CYC for AAV (RAVE[23] and RITUXVAS[24]). No evidence was identified comparing RTX with other relevant comparators such as AZA, MMF or MTX. Non-RCT evidence satisfying the inclusion criteria was identified and tabulated. Relevant data from the published articles of the RAVE and RITUXVAS trials were extracted and presented in full. Both trials were critically appraised using standard superiority RCT criteria. The findings of RAVE and RITUXVAS were reported, but no synthesis was performed. In the absence of head-to-head trials comparing RTX with other comparators listed in the scope and decision problem, a search should have been conducted to identify relevant RCTs for an indirect comparison.

The MS includes a systematic review of the clinical effectiveness evidence. However, the review is inadequately reported in parts. The inclusion and exclusion criteria are not consistent with the NICE scope and final decision problem; the limited scope of the submission was justified by the manufacturer based on the limitations of the anticipated licence. The searches were inadequate and poorly reported, but this was corrected by the manufacturers within their response to the ERG's clarification questions.[20] Methods of study selection and data extraction were not reported. The choice of critical appraisal checklist and criteria is not described or justified and, given that RAVE was a non-inferiority trial, additional criteria specific to this type of study design should have been considered. The Results section of the MS consisted simply of a description or overview of each RCT and its results (see MS, Section 6.5) or the reproduction of the abstracts of the non-RCTs (see MS, Section 6.7). There was no meta-analysis or narrative synthesis. The RITUXVAS trial is presented as supporting evidence only on account of the intervention being RTX + CYC and the population being generalised, "severe" AAV with renal involvement; RAVE is considered the only "pivotal" trial (see MS, Section 6.2.5). However, clinical advice and the scope and decision problem criteria suggest that

RITUXVAS is equally relevant, especially as the choice of comparator (IV CYC) arguably represents the more common comparator in expert clinical practice in the NHS. The two trials should therefore be given equal consideration in any synthesis. The question whether RTX or RTX+CYC offers the optimum intervention is also equivocal,[4,25] so the RITUXVAS data are also pivotal. The submission presents the results of subgroup analyses, but there is no published, dateable evidence that these were pre-specified and a number of these analyses do not relate directly to the outcomes in the decision problem.

The submission provides a separate summary of the safety profile of RTX for AAV (see MS, Section 6.8). This was neither conducted nor reported as a systematic review. However, this section should have either formed part of the efficacy systematic review or should have been a separate, adverse events systematic review.

No details were given of the inclusion or exclusion criteria, or the study selection or data extraction processes for the adverse events review (see MS, Section 6.8). No critical appraisal of the evidence was conducted beyond the assessments of the RAVE and RITUXVAS trials undertaken for the efficacy review (see MS Sections 6.4 and 10.3). Results consisted only of a simple description or overview of the RAVE and RITUXVAS trial evidence. Brief characteristics of other studies identified as relevant were tabulated (see MS, Table 33). There was no attempt to meta-analyse or narratively synthesise this evidence.

4.1.1 Searches and study selection

Searches

The manufacturer's search methods and electronic strategies were appraised by the ERG group. The ERG sought clarification from the manufacturer where weaknesses in the manufacturer's search methods were considered to impact the performance of the search and may thus potentially lead to the omission of studies.

The manufacturer's searches for direct clinical evidence were adequately reported and strategies were explained. The manufacturer clearly acknowledges CYC as the most relevant comparator in the induction of remission; however, separate searches were not conducted by the manufacturer for other evidence that could have been used to inform indirect comparisons against other drugs. Separate adverse events searches were not conducted for either RTX or CYC.

The sensitivity of the search strategies in the submission was open to question due to i) omission of free-text synonyms for "rtx" in all search strategies i.e. "rituximab" or "mabthera", "rituxan" and "rituxin". ii) omission of the subject heading "Vasculitis". These limitations were raised in the ERG

clarification letter and the manufacturer re-ran searches and produced responses which are described below.

The manufacturer searched the minimum required databases. Searches for ongoing or completed and unpublished trials – using sources such as ClinicalTrials.gov, metaRegister of Controlled Trials, and the WHO International Clinical Trials Registry Platform - were not reported. In addition, the ERG suggests that the manufacturer should have searched relevant society websites such as the European Vasculitis Society (EUVAS) and specialist conference abstracts, such as those associated with the International Vasculitis & ANCA Workshop.

The ERG considered that the language restriction to English only publications was too restrictive. It is not clear whether relevant foreign language publications may have been missed.

Translation of the strategies from Medline and Embase to the Cochrane Library was not consistently applied; intervention terms were omitted in the Cochrane Library search. The manufacturer acknowledged and rectified this in their Clarification Response letter.[20]

Four search priority requests were made by the ERG to the manufacturer in the clarification letter:

1) To review and conduct searches for direct evidence using synonyms such as rituximab, "Rituxan" or "Mabthera" and including a broader subject heading "Vasculitis". Revision and re-run of searches by the manufacturer identified a further 35 records but the manufacturer stated that these were all related to trials that had already been found by the initial search. Hence the manufacturer stated that no new publications were identified that were relevant for consideration within this appraisal (see manufacturer's Clarification Response, p.25).

2) To search within clinical trials registers for completed and unpublished trials. The manufacturer reported that they had searched ClinicalTrials.gov and WHO ICTRP, and found 17 studies. However, the manufacturer reported that with the exception of RAVE and RITUXVAS, these studies were found to be either unpublished, not involving the licence population, not yet initiated, or, in one case, the study was withdrawn prior to enrolment.

3) To carry out separate adverse events searches for both RTX and CYC. According to the manufacturer, safety information for RTX were acquired from the US and EU regulatory dossier (see p.3 and p.10 of manufacturer's Clarification Response[20]). The ERG noted that a direct search for adverse event data would have been preferable and provided suggested search strategies to the manufacturer. The manufacturer ran these search strategies and identified 2,284 papers on RTX and 8,485 on CYC. Given the large number of records to sift

and the short time scales, the number of studies relevant to the decision problem was not determined by the manufacturer.

4) To carry out indirect comparator searches for CYC compared to other drugs that would be used in sequence such as MTX or MMF. The manufacturer did not conduct separate indirect comparison searches and re-stated their belief that the evidence base for the additional comparators is not within the scope of the appraisal as they do not reflect the population concerned. The ERG designed and carried out searches for trials comparing CYC and MMF (see Appendix 1 for ERG search strategies and a summary table of ERG searches) in combination with a sensitive RCT filter and identified 715 records. The ERG found two studies that appeared relevant for the decision problem set out for this appraisal. In their clarification response the manufacturer acknowledged the MYCYC trial, comparing MMF to CYC, but stated that the results of this study were not yet published.[20] However, the ERG searched and found one published conference abstract reporting early results from this trial.[22]

Study selection

The process of study selection was neither described nor evaluated within the MS (e.g. citations screened independently by more than one reviewer). The PRISMA diagram presented in the original submission was inadequate as a record of the search and selection process, but this has been updated adequately in response to a request of the ERG (see manufacturer's Clarification Response Appendix 1; although the updated PRISMA diagram contains an error on numbers of full papers excluded, i.e. 709 for 706).[20] Details of all studies excluded at full paper stage were provided and the reason for exclusion was given (see manufacturer's Clarification Response: Appendix 2, Table 9).[20]

4.1.2 Inclusion criteria

The reported criteria for the effectiveness review are detailed in Box 1:

Box 1: Inclusion and exclusion criteria reported by the manufacturer

Inclusion criteria Published papers or abstracts which evaluated the following:	
RTX had to be the major focus of the paper, in order to eliminate papers which mentioned RTX as part of a discussion of treatments for rheumatoid arthritis AAV had to be a major focus of the paper, in order to eliminate papers covering the use of RTX in other autoimmune diseases Patient population should consist of those patients who were receiving RTX for induction of remission (or treatment of flare), i.e. not maintenance data, to be consistent with the proposed RTX licence Correct dosage of RTX 375mg/m^2 body surface area once weekly for 4 weeks Clinical trial data Documents relating to humans	
Exclusion criteria Published papers or abstracts which evaluated the following were excluded:	**Rationale / justification***
Any papers providing a review, update or commentary on data published elsewhere Any papers which only mentioned RTX within a discussion of treatments for AAV or other auto-immune diseases Animal studies or in vitro research	To ensure no duplication of results / data No data in these papers Only human data relevant to decision problem
Case reports Studies where there were data for fewer than 20 patients Posthoc subgroup analyses	Not statistically robust analyses
Papers covering Churg-Strauss syndrome Paediatric studies Incorrect dosage of RTX Maintenance of remission only	Not in the licence, i.e. induction of remission only using 4x375mg/m^2 dose of RTX for adults with generalised, “severe” AAV only

**provided in manufacturer’s Clarification Response: Appendix 3 Table 10*

The inclusion criteria are generally in line with the scope and the decision problem, but there are a number of variations, which are explained with reference to the licensed dose of RTX for this indication, i.e. induction of remission only in adult populations with generalised, “severe” AAV only, and for the licensed the 4x375mg/m^2 regimen only.

Population: As noted in Section 3.1 the NICE scope considers all patients with AAV (mild and severe), but the decision problem outlined by the manufacturer restricts evidence to induction of remission in generalised, "severe" AAV adult populations only. This is reflected in the inclusion criteria used in the manufacturer's search. Similarly, papers investigating maintenance therapy were excluded from the effectiveness review.

Intervention(s): As noted in Section 2.2, a dose of 2x1g of RTX is more commonly used than 4x375mg/m^2 in the NHS. The effectiveness review restricts itself to evidence for the 4x375mg/m^2 regimen only; Evidence is not sourced on the 2x1g dose, which is described variously as the "incorrect dose" and "off-label". While the ERG accepts that the 4x375mg/m^2 regimen represents the dose that was submitted for licensing, it is necessary to point out that the scope and decision problem covered both doses and regimens and that the relative efficacy and safety of the two is uncertain. This point has also been made in the literature.[4,25] This is important as the 4x375mg/m^2 regimen represents the higher overall dose and this might be considered an issue given the dose reduction principles that govern the rationale behind the selection and evaluation of the alternative immunosuppressive therapies for this population.[4,25] This is acknowledged in the MS (see MS, Section 2.6, p.26-27), which notes that, "The 1,000 mg dose of RTX is widely-used off-label for AAV and gives less exposure compared with 375mg/m^2 once weekly for 4 weeks schedule ...[and] we believe that the 2 x 1,000 mg dose has widespread use across the UK."

Comparators: As noted in Section 3.3, the NICE scope and decision problem requires the consideration of RTX versus all possible comparators for this indication, i.e. MTX, AZA and MMF, as well as CYC, but the effectiveness review restricts itself to evidence for CYC on the basis that this is the "standard remission induction agent" (See MS, Sections 2.4 and 5) for severe AAV populations who would otherwise only be prescribed CYC. However, there are RCTs of potential relevance to an indirect comparison, which compare MMF with CYC for induction of remission in a generalised, "severe" AAV population who might otherwise be prescribed CYC (CYC is the therapy in the control arm) (see Section 4.5 below). These trials were not sourced or included (and no indirect comparison undertaken) because of the restriction of the clinical effectiveness review to the "only relevant comparator" of CYC only (see manufacturer's Clarification Response: Priority requests 13 and 22[20]). The potential relevance of the NORAM RCT[26,27] comparing MTX and CYC in a less severe population was acknowledged by the manufacturer, but dismissed on clinical advice (see manufacturer's Clarification Response: Priority request 4[20]). The manufacturer stated they were aware of the ongoing trial NCT00414128 or EUDRACT: 2006-001663-33 (MYCYC: MMF vs CYC) (see manufacturer's Clarification Response Priority request 4[20]).

Outcomes: All outcomes specified in the scope and decision problem, except duration of remission, are considered.

Summary

The manufacturer's decision to restrict the review to the more limited population and intervention specified in the licence application means that the effectiveness review does not wholly satisfy the scope or decision problem specified by NICE. However, with the exception of some trial evidence of relevance for an indirect comparison, the ERG has not identified any additional relevant head-to-head RCT evidence comparing RTX with any of the named comparators for induction or maintenance of remission in any AAV population.

4.1.3 Critique of data extraction

All key efficacy and safety data appear to have been extracted accurately from the two principal studies, however the process of data extraction was neither described nor evaluated within the MS (e.g. using double data extraction or verification). Only the most basic study characteristics were extracted from the identified non-RCT efficacy studies. Consequently, the submission made no use of data from studies other than RAVE and RITUXVAS in the analysis or to illuminate or facilitate the interpretation of the efficacy or safety data from these two principal trials.

4.1.4 Quality assessment

The quality assessment consisted of the critical appraisal of the RAVE and RITUXVAS trials using standard RCT criteria (see MS[8] Sections 6.4 and 10.3). RAVE was both a superiority and non-inferiority trial, depending on analysis and outcome, but only criteria assessing superiority were applied. This was corrected at the request of the ERG (see manufacturer's Clarification Response Appendix 6, Table 12[20]). RITUXVAS was a superiority trial and was appraised as such. The submission found that the RCTs had a low risk of bias across all criteria. The submission did not report a high risk of bias on any criteria, but considered criteria on blinding to be "Not applicable" to the RITUXVAS trial as it was "open-label". This is not an appropriate judgement because the open-label nature of the trial renders it at high risk of performance and detection bias, i.e. patients and outcome assessors are aware of the treatment received and their judgments might be altered as a result (see Appendix 2).

The ERG applied a combination of the Cochrane risk of bias tool to appraise risk of bias within each trial,[2] as well as the non-inferiority trial extension of the CONSORT statement for the RAVE trial.[3] The findings of this appraisal are provided in Appendix 2 and were generally consistent with the reported assessment in the MS, except for the high risk of performance and detection bias in the RITUXVAS trial as noted above. The RAVE trial appears to be at low risk of bias across all domains,

but a number of issues should be noted. First, the source of the criterion of non-inferiority for the primary efficacy endpoint is neither explained nor justified in the RAVE publication or supplementary files, but only in the MS. The 2005 WGET trial appears to be acting as the reference for the primary endpoint and non-inferiority margins.[28] However, this RCT was conducted in much younger patients (mean age of 47.5 years in the CYC arm) with generalised AAV; patients with GPA only (not MPA) and with a primary endpoint of sustained remission (BVAS/WG of 0 for at least 6 months). The rate of sustained remission in the CYC arm was reported as 75% not 70%. The trial reported by Jayne *et al*.[12] by contrast considered older patients (mean age 59) with severe AAV GPA and MPA and ANCA-positivity (ANCA-negative patients were eligible if there was histologic confirmation of vasculitis); a comparable dose of oral CYC (2mg/kg) and a definition of remission which required only a disease score of 0 on BVAS. The trial achieved a clinical remission rate of 93%. This trial arguably would have acted as the better reference trial for non-inferiority for the RAVE trial. There exist some questions therefore concerning the appropriateness of the reference trial for determining the non-inferiority criteria. Second, two protocols are available for RAVE, one of which only exists in a version that post-dates the inception of the trial (2009) and is only available as an online supplementary file with the 2010 publication (RAVE[23]). This latter protocol is the only source of information on randomisation and allocation concealment. The primary endpoint was updated during the course of the RAVE trial from complete remission to complete remission with prednisone taper. While the updated endpoint represents a stricter definition with potentially lower rates of remission than reported elsewhere, it is also possible that the results for this updated primary endpoint are relatively more favourable to RTX than the original endpoint. For example if one considers only complete remission with a prednisone taper threshold <10mg/d then the difference between RTX and the comparator is slightly smaller (i.e. 70.7% vs 62.2%: rather than 63.6% vs 53.1% see Table 2 below, with a difference in p value of $p<0.001$ and $p=0.1$, see 4.2 below). Subgroup analyses are described as "pre-specified" in both the RAVE publication and the MS but details of these analyses are not provided in any of the available published protocols. Rather they are given only in the "Statistical Analysis Plan", an extract of which was provided by the manufacturer in response to a request by the ERG (see manufacturer's Clarification Response: Priority request 12[20]), which indicated that *"the consistency of the primary endpoint results* ***was*** *explored"*, suggesting a *post hoc*, rather than *a priori* approach. Both trials might therefore be considered as being at low risk of bias, but such an assessment must take into account the caveats described.

4.1.5 Evidence synthesis

The submission did not include an evidence synthesis. The principal RCTs were not combined in a meta-analysis because of clinical heterogeneity (slightly different populations, different interventions and different regimens for the comparators). The manufacturer did not perform a narrative synthesis either. The Results section of the MS consisted of the reproduction and description of the published

findings of the trials, including the presentation of multiple published sub-group analyses. Despite the presentation of non-RCT evidence and studies of adverse events, no use was made of any evidence other than from the two key trials. The manufacturer justified this by claiming that the data were "problematic for providing robust conclusions" (see manufacturer's Clarification Response: Priority request 15 Priority request 4[20]).

The choice not to combine RAVE and RITUXVAS in a meta-analysis was justified, although these trials have been pooled in a meta-analysis elsewhere.[29] However, a formal narrative synthesis should have been performed, comparing and contrasting the findings of the two trials and exploring and explaining their similarities and differences. Subgroup analyses are described as "pre-specified" in the MS and the published paper,[23] but details are not provided in any of the available protocols for RAVE.

In the absence of relevant head-to-head trial evidence for RTX against all comparators in the scope, i.e. CYC, MTX and MMF, and the availability of relevant RCTs comparing CYC with MMF for patients with generalised, "severe" AAV, then there is an argument that an indirect comparison should have been performed and presented (See Sections 4.2 and 4.5 below).

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

The submission identified two relevant RCTs: RAVE and RITUXVAS, and seven non-RCTs. The ERG report will focus on the two principal trials. Details relating to the lack of sufficient data for use from the non-RCTs are given in the manufacturer's Clarification Response: Priority request 9 and Appendix 4, Table 11.[20] The submission presents the RAVE trial as the most relevant because it evaluates RTX alone (as per the anticipated licence) in a generalised, "severe" AAV population vs oral CYC rather than RTX plus 2-3 pulses of IV CYC vs IV CVC in a severe AAV population with renal involvement. However, clinical advice received by the ERG suggests that both RCTs are clinically important and that RTX+CYC should equally be considered as the efficacy and safety of 4x375mg/m^2 RTX alone has not been demonstrated relative to 4x375mg/m^2 RTX plus reduced dose CYC. This issue is acknowledged in the literature[4,25] and is also accommodated by the scope and decision problem.

Details of the RAVE and RITUXVAS trials are provided in Table 2.

Table 2: Study characteristics

Trial name	Citation	Inclusion criteria	Exclusion criteria	Phase	Intervention	Comparator	Follow-up
RAVE[23]	Stone 2010[23], Stone 2009[30], Specks[31] , Stone[32]	New diagnosis or relapsing AAV (GPA or MPA according to Chapel Hill consensus conference definitions); ANCA positivity; BVAS/WG of ≥3 (only 66% have renal involvement, Geetha[33] = 61% vs 63% remission – like RITUXVAS); severe disease, i.e. they would normally be treated with CYC; Aged 15 years or older; at least 40kg weight	CSS; limited disease activity that would not normally be treated with CYC; mechanical ventilation because of alveolar hemorrhage; creatinine level >4.0mg/dl attributed to renal failure from a current flare; receipt of oral or IV CYC within 4 months prior to enrolment; history of adverse effects from standard therapy (i.e. intolerant to CYC); GC for no longer than 14 days before screening	Induction de novo and post-relapse on CYC	RTX 375mg/m^2 plus daily CYC placebo, plus AZA placebo if achieve remission 3-6 months, plus PD (1mg/kg/d [max.80mg/d], reduced to 40mg/d no later than end of week 4, then stepwise reduction of dose every 2 weeks, tapered so that discontinued by 6 months if have remission and no disease flares)	CYC oral 2mg/kg/d, adjusted for renal insufficiency, plus 2mg/kg/d AZA if achieve remission 3-6 months, plus PD (1mg/kg/d tapered so that discontinued by 6 months if have remission and no disease flares)	6, 12 (Specks[31]) and 18 months (Stone[32])
RITUXVAS [24]	Jones 2010[24], Jones[34-36]	New diagnosis AAV; ANCA positivity; renal involvement (≥30 red cells per high power field on urinanalysis)	Previous CYC, (greater than 2 weeks of an oral or IV pulse CYC regimen); Co-existence of another multisystem autoimmune disease, e.g. SLE, CSS; Hepatitis B e antigen positive or Hepatitis C antibody positive; Known HIV positive (HIV testing will not be a requirement for this trial); Previous malignancy (usually exclude unless agreed with trial co-ordinator).	Induction	RTX 375mg/m^2 x 4 plus with 1st and 3rd RTX infusions (plus 3rd dose if there was progressive disease within first 6 months), plus PD (1mg/kg/d reduced to 5mg/d by 6 months)	CYC IV pulse 15mg/kg for 3-6 months, followed by AZA if achieve remission, plus PD (1mg/kg/d reduced to 5mg/d by 6 months)	12 months

CSS: Churg-Strauss syndrome: AAV: ANCA-associated vasculitides (AAV): ANCA: Anti-neutrophil cytoplasmic antibody: CYC: cyclophosamide; AZA: azathioprine; RTX: rituximab; PD: prednisolone; GC: Glucocorticoids; SLE: systemic lupus erythematosus

Table 3: Study outcomes

Trial name	Citation	Primary endpoint(s)	Secondary efficacy endpoints	Adverse events*
RAVE[23]	Stone 2010[23], Stone 2009[30], Specks[31], Stone[32]	▪ Remission: absence of disease activity, i.e. BVAS/WG of 0 with completion of prednisolone taper by 6 months	▪ Remission: absence of disease activity, i.e. BVAS/WG of 0 with PD dose of <10mg/d by 6 months ▪ Cumulative GC dose ▪ Rates of disease flares: an increase in BVAS/WG of 1 point or more ▪ SF-36	▪ Malignant conditions ▪ Leucopenia or thrombocytopenia of grade 2 or higher ▪ Infections (grade 3 or higher) ▪ Death (all causes) ▪ Drug-induced cystitis ▪ VTE events ▪ Stroke ▪ Hospitalisations ▪ Infusion reactions
RITUXVAS[24]	Jones 2010[24], Jones[34-36]	▪ Sustained remission: absence of disease activity, i.e. BVAS of 0 for at least 6 months; ▪ Relapse: recurrence or new appearance of any disease activity, as reflected by BVAS, attributable to vasculitis ▪ Rates of serious adverse events**	▪ Time to remission, i.e. absence of disease activity, i.e. BVAS of 0 for 2 months; ▪ Change in BVAS between 0-3 months ▪ Change in the GFR ▪ Prednisolone dose ▪ SF-36 from baseline to 12 months ▪ VDI from baseline to 12 months	**Serious adverse events, including hospitalisations and cancer ▪ Infections ▪ Death ▪ Infusion reactions ▪ Hematologic events: anaemia, neutropenia, thrombocytopenia, hypoglammaglobulinemia ▪ Graded events 1-2 and 3-5

**All graded by the National Cancer Institute's Common Terminology Criteria (http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm)*

BVAS/WG: Birmingham Vasculitis Activity Score/Wegener's granulomatosis; PD: prednisolone; GC: Glucocorticoids; GFR: Glomerular filtration rate ; VDI Vasculitis Damage Index; SF-36: Short Form-36; VTE: venous thromboembolism

4.2.1 Results

RAVE: Primary and secondary efficacy outcomes

RTX satisfied the non-inferiority criteria compared with oral CYC for both complete remission with PD taper ($p<0.001$) or for complete remission for a PD dose of <10mg at 6 months ($p = 0.10$) (see Stone *et al.*[23]) In the intention to treat (ITT) population 63.6% (95% CI 54.1%, 73.2%) in the RTX group and 53.1% (95% CI 43.1%, 63.0%) in the CYC group achieved complete remission at 6 months, and 70.7% (95% CI 61.7%, 79.7%) in the RTX group and 62.2% (95% CI 52.6%, 71.8%) in the CYC group achieved remission with a PD dose of <10mg at 6 months (see Table 4). There was no significant difference between arms for rates of severe flares per patient month (0.011 for RTX compared to 0.019 for CYC, $p = 0.29$) or the rates of limited disease flares per patient month (0.026 in both groups, $p = 0.98$). There were also no differences in Vasculitis Damage Index (VDI) or Quality of Life (SF-36) scores. RTX appears to be superior to oral CYC for patients with relapsing disease at baseline for the exploratory endpoint of complete remission (BVAS of 0, and without reference to PD dose), ($p = 0.02$). Exploratory subgroup analysis of complete remission in newly diagnosed patients gave an absolute difference of -4.2% (95% CI -23.6%, 15.3%, $p = 0.673$) in favour of CYC (60.4% in the RTX group compared to 64.6% in the CYC group). A similar analysis in patients with recurrent disease gave an absolute difference of 24.7% (95% CI 5.8%, 43.6%, $p = 0.013$) in favour of RTX (66.7% in the RTX group compared to 42.0% in the CYC group).

It is noteworthy that the data presented on flares in the MS differ slightly from those presented by Stone *et al.* (2010),[23] seemingly as the result of minor differences in the number of limited and severe flares attributed to each treatment group. In Stone *et al.* (2010) the rate of severe flares per patient-month are stated to be 0.011 and 0.018 in the RTX and CYC groups, respectively ($p = 0.30$), and the rate of limited flares per patient-month are stated to be 0.023 and 0.027 in the RTX and CYC groups, respectively ($p = 0.81$).[23]

Table 4 provides a summary of the primary and secondary efficacy outcomes in RAVE (including ITT Population).

Table 4: (reproduced and adapted from MS)

	RTX (n=99)	CYC (n=98)	Absolute Difference (95% CI)	p-value [a]
Primary endpoint				
n analysed [b]	98	95		
% in complete remission [c] (95% CI)	64.3 (54.8, 73.8)	54.7 (44.7, 64.8)	9.5 (-4.3, 23.4) [d]	0.177
ITT population				
n analysed [b]	99	98		
% in complete remission [c] (95% CI)	63.6 (54.1, 73.2)	53.1 (43.1, 63.0)	10.6 (-3.2, 24.3)	0.132
Secondary and tertiary endpoints				
% in remission on < 10 mg/day of prednisone (95% CI)	70.7 (61.7, 79.7)	62.2 (52.6, 71.8)	8.5 (-4.7, 21.6)	0.208
% with severe flare	5.1	10.2		
Rate of severe flares per patient-month	0.011	0.019*		0.293*
% with limited flare	12.1	14.3		
Rate of limited flares per patient-month	0.026*	0.026*		0.98*
Mean (SD) BVAS/WG AUC over first 6 months	1.29 (1.33)	1.25 (1.03)		
Median cumulative prednisone dose (1,000mgs) (95% CI)	3.3 (1.0, 6.9)	3.5 (0.7, 8.3)		0.055
Mean change from baseline in ESR (95% CI)	-14.4 (-18.7, -10.1)	-9.3 (-15.6, -3.0)	7.6 (2.2, 13.1)[e]	0.006
Mean change from baseline in CRP (95% CI)	-2.69 (-5.44, 0.06)	-2.84 (-7.07, 1.40)	0.61 (-0.5, 1.73)[e]	0.278
Exploratory endpoints**				
Patients (%) in remission (BVAS/WG = 0) (95% CI)	80.8 (73.0, 88.6)	66.3 (60.0, 75.7)	14.5 (2.3, 26.7)	.021
% of newly diagnosed pts in complete remission	60.4	64.6	-4.2 (-23.6, 15.3)	0.673
% of relapsing pts in complete remission	66.7	42.0	24.7 (5.8, 43.6)	0.013

AUC = area under the curve; BVAS = Birmingham Vasculitis Activity Score; CI = confidence interval; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; WG = Wegener's granulomatosis; WCI = worst case imputation; a, p-value is from a chi-square test of superiority; b, Patients with non-missing results for complete remission at 6 months after randomisation; c, Defined as BVAS/WG = 0 with successful completion of glucocorticoid taper at Month 6; d, The lower limit of the 95% CI for the absolute difference, -4.3%, was greater than -20% and thus met the protocol-specified non-inferiority criterion; e, the differences in ESR and CRP are adjusted by ANCA status, site, and baseline value.

**Figures that do not match those presented in Stone et al. (2010)* [23]

RITUXVAS: Primary and secondary efficacy outcomes

Sustained remission (primary endpoint) occurred in 25 of 33 patients in the RTX group (76%) and 9 of 11 patients in the control group (82%). The absolute difference in sustained remission with RTX as compared with CYC was −6 percentage points (95% CI, −33 to 21; p = 0.68). Among the patients who survived, 93% of the patients in the RTX group and 90% of the patients in the control group had sustained remission (p = 0.80). Table 5 provides a summary of the primary efficacy outcomes in RITUXVAS.

Table 5: (reproduced and adapted from MS)

	RTX + low dose CYC n (%)	CYC n (%)
Achieved sustained remission at 12 months	25 (76)	9 (82)
Reasons for non-response		
Death	6 (18)	1 (9)*
Relapse within 6 months after remission	1 (4)	1 (9)
Re-treated for incomplete remission	1 (4)	0

** Another patient died at 19 months, for a total of 2/11 (18%) deaths in the control arm*

There were no significant differences between arms for any of the secondary efficacy outcomes in terms of median time to remission (p = 0.87); median BVAS and prednisone doses (see Figure 3 in Jones[24]); median estimated GFR (p = 0.14 for the comparison of medians), median change in the VDI (p = 0.38) or SF-36 (p = 0.36). The CYC group had an improved mental composite SF-36 score in comparison with the RTX + low-dose CYC group (p = 0.04) but this difference was largely accounted for by 2 patients' scores in the RTX + low-dose CYC group. Exclusion of these 2 patients' data resulted in a p-value of 0.32.

4.2.2 Adverse events

A brief summary table of the adverse events common to both trials is given in Table 6.

Table 6: Number of adverse events in any arm

Adverse events	RAVE (6 months follow-up)	RITUXVAS (12 months follow-up)
	RTX (n=99) / CYC(n=98)	RTX+CYC (n=33) / CYC (n=11)
Serious adverse events (≥grade 3)	31/12	73 / 85
Deaths	1 / 2	6 / 2 (both 18%)
Cancer (patients)	1 / 1†	2 / 0
All infections	NR	19/ 7
Serious infections (≥grade 3)	7 / 7	7 / 3
Thrombocytopenia	3 / 1*	1 / 0
Neutropenia	NR	2/ 1
Leucopenia (≥grade 2)	3 / 10	NR

**(≥grade 3) †5/2 for >6 months follow-up (see text)*

RAVE

There were no reported significant differences between the treatment groups in almost all adverse events outcomes, but some notable disparities regarding leucopenia and malignancies. More patients in the control group than in the RTX group had one or more of the predefined selected adverse events: 32 (33%) versus 22 (22%) (p = 0.01), but more episodes of Grade 2 or higher leucopenia in the control group (10 vs. 3) accounted for most of this difference. Malignant conditions developed in 7 patients after 6 months, i.e. 6 of 124 (5%) in the RTX arm, as compared with 1 of 73 patients without exposure to RTX (1%, p = 0.26).

RITUXVAS

A total of 31 severe adverse events occurred in 14 of the 33 patients in the RTX group (42%) and 12 severe adverse events occurred in 4 of the 11 patients in the control group (36%). There was no significant difference between groups for incidence rates for severe adverse events (p = 0.77). The safety profile of the two regimens appears comparable; only infections and mortality appear potentially higher in the RTX arm. Six of the 33 patients in the RTX group (18%) and 2 of the 11 patients in the control group (18%) died (p = 1.00). However, the causes of death were infections (in 3 patients in the RTX group and in 1 patient in the control group), cardiovascular disease (in 1 patient in the RTX group and in 1 patient in the control group), and complications of end-stage renal failure (in 2 patients in the RTX group).

4.2.3 Comments

Population: The populations were limited in the two RCTs to those with generalised, "severe" AAV and ANCA positivity. RAVE considered both newly diagnosed and relapsing patients; RITUXVAS considered only the former. Age and renal insufficiency are known predictors of adverse events for AAV[16,37] and the population in the RAVE trial was comparatively young (mean / standard deviation = 54.0±16.8 vs 51.5±14.1) and had a creatinine of <4.0mg/dl. This suggests that the rates of adverse events reported by RAVE might be lower than one would expect for the AAV population presenting in practice. By contrast the median age of the participants in the arms of the RITUXVAS trial was 67-68 years and severe AAV with renal involvement were inclusion criteria.

Interventions: Both trials used large doses of prednisone according to expert commentary[4,5,38] and neither used the commonly prescribed 2x1g dose of RTX, hence there is uncertainty regarding the relative efficacy of these two regimens.

Comparators: The similarity in rates of adverse events is potentially unexpected in the two RCTs (RITUXVAS hypothesised superiority of RTX for these outcomes), especially for the RAVE trial

which compared RTX with oral CYC, a comparator with a potentially worse safety profile than the IV pulse CYC used in the RITUXVAS trial.[39-41]

Outcomes: Both trials used BVAS/WG to evaluate remission (a possible subjective element to this measure has been noted in the literature,[42,43] though the scale has been found to have good inter-observer agreement for GPA populations.[44] It is noteworthy that BVAS/WG, as opposed to BVAS, was developed specifically for a GPA population, whereas both RAVE and RITUXVAS included GPA and MPA patients. The definition of remission used in the two trials is not the same. The RAVE primary and secondary efficacy endpoints regarding remission are more restrictive than are often used, requiring the satisfaction of prednisone dose criteria. However, results are also reported for this trial using a more conventional definition of complete remission (BVAS/WG of 0), without any prednisone dose criteria (see MS,[8] Table 15). The findings for these different definitions are not statistically significantly different. Only the subgroup of relapsing patients (who have received one or more prior doses of CYC, MTX or AZA) demonstrates a significant difference in the primary endpoint, with RTX being superior to CYC (p=0.01).

Adverse events: Neither trial demonstrated superiority of the RTX or RTX+reduced dose CYC regimens compared with CYC. This is of note given that part of the rationale behind the use of RTX is its relatively better safety profile compared to CYC. It is true that the short duration of the trials might not demonstrate the relative safety of the regimens for an outcome such as fertility,[23] but this is not the case for other adverse events. The authors of the RAVE trial claim that leucopenia is likely to be higher in practice in patients receiving CYC than observed in the trial[23] but present no evidence to support this. Rates of infections appear to be higher in the RTX arm in the RITUXVAS trial, and rates of malignancies are higher in the RTX arm of the RAVE trial. Some concerns have been expressed in the literature regarding these findings,[4,5] but the trial was only of short duration and the authors present possible explanations for these relatively higher rates in the RTX arms (e.g. malignancies still being within an expected range), and also potentially attributable to other treatments, such as CYC and AZA, to which patients were exposed (RAVE Supplementary file[23]).

No additional relevant RCTs comparing RTX vs CYC were found by the ERG. No RCTs were identified by the MS or ERG comparing RTX with the other comparators outlined in the scope and decision problem for this population and indication (i.e. AZA, MTX, MMF). However, three published RCTs and two ongoing RCTs were identified by the ERG comparing induction regimens of CYC vs MMF or MTX in moderate or severe AAV populations (see Section 4.5 for more details on these trials). Together with the RTX vs CYC trials (RAVE and RITUXVAS), this trial evidence might enable an indirect comparison/Mixed Treatment Comparison (MTC) to evaluate the relative efficacy of RTX and the comparator of MMF, which is listed in the NICE scope and decision problem (see Section 4.5 below). All studies were identified from the search conducted by the ERG information

specialist (RW) (see Section 4.5 below). The submission did not provide any reason for the exclusion of these studies, other than that it chose only to consider trials of RTX against the "only relevant comparator" (see manufacturer's Clarification Response priority request 13[20]), i.e. CYC, for adult populations with generalised, "severe" AAV. Thus the other comparators listed in the scope were excluded.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

No indirect comparison or MTC was conducted, though such an analysis was possible – see Section 4.5 below.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

No indirect comparison or MTC was conducted, though such an analysis was possible – see Section 4.5 below

4.5 Additional work on clinical effectiveness undertaken by the ERG

The MS excluded comparators other than CYC, but the following additional comparators were covered by the scope and decision problem: MTX, AZA and MMF. The ERG therefore conducted a search (as specified in Appendix 1) to identify any relevant published and unpublished RCT evidence using these treatments in AAV populations.

The search and study selection process identified five published or ongoing RCTs using these treatments for induction of remission in AAV populations designated by the scope and decision problem, and similar to those in the RAVE and RITUXVAS trials.[26,45-48] These trials are summarised in Table 7 below. The potential relevance of the NORAM RCT[26,27] comparing MTX and CYC was acknowledged by the manufacturer, but dismissed on clinical advice (see manufacturer's Clarification Response: Priority request 4[20]). However, the NORAM trial does appear to satisfy the decision problem and scope. The manufacturer stated that they were aware of the ongoing trial MYCYC, NCT00414128 or EUDRACT: 2006-001663-33 (MMF vs CYC) (see manufacturer's Clarification Response: Priority request 4[20]).

Table 7: Possible RCTs for indirect comparison with CYC as comparator

Citation	Population	Phase	Number	Intervention/dose	Comparator/dose	Treatment duration
Chen 2012[45]	ANCA positive AAV, "Generalized"[non-organ threatening AAV]	Induction *de novo*	29	CYC IV pulse (0.5-0.75g/m^2/month)	MMF (2g/d or 1.5g/d if <50kg)	6 months
Hu 2008[46]	Severe AAV , 87.5% ANCA positive; 18+ years and with renal involvement. Excludes very severe disease, i.e. creatinine of >500 or life threatening lung haemorrhage or CNS disease	Induction *de novo*	35	CYC IV pulse (0.75-1g/m^2/month)	MMF was given as 2.0 or 1.5 g/day (for body weight <50 kg) for 6 months.	6 months
De Groot 2005[26], Faurschou 2012[27] (NORAM)	Generalized non-organ threatening severe AAV. Excludes creatinine of >150 or life/organ threatening diseas	Induction *de novo* and maintenance	100	CYC oral 2mg/d	MTX 20-25mg/w	18 months, FU 6 years

Table 8: Relevant ongoing indirect comparison RCTs

Trial #	Population	Phase	Intervention/dose	Comparator/dose
NCT00414128 MYCYC[48]	New diagnosis of AAV; ANCA positivity; BVAS ≥3. Excludes life threatening vasculitis (lung haemorrhage, GI perforation, cerebral / cardiac vasculitis) or renal disease and GFR fall >20% in < 2 weeks	Induction *de novo*	CYC (NR)	MMF (NR)
NCT00103792[47]	Relapsed AAV. Excluded patients: "severe lung haemorrhage" or creatinine >500	Induction post-relapse (1st or 2nd)	CYC oral 2mg/d	MMF 2g/d

The MYCYC trial population is similar to the RAVE trial, while the Hu 2008 and NCT00103792 trials cover populations with a slightly more "severe" form of AAV than the RAVE trial, and Chen 2012 and NORAM cover populations with a slightly less "severe" form of AAV. The ERG contends that, using this evidence, an indirect comparison might be possible to determine the relative efficacy and safety of RTX compared with MMF and MTX (albeit for a slightly different population in the latter case). Such an analysis could have been conducted and submitted by the manufacturer given the required scope of the submission and the decision problem. The relative efficacy and safety of RTX (and CYC) compared with alternative treatments for induction of remission in generalised, "severe" AAV populations is therefore uncertain.

4.6 Conclusions of the clinical effectiveness section

The MS identified the two RCTs comparing RTX with CYC as induction therapy for adults with generalised, "severe" AAV. There is no head-to-head RCT which directly compares RTX with other relevant comparators, such as AZA, MTX or MMF for severe AAV. The submission did not conduct a meta-analysis or synthesis and thus reports the results (equivalence or superiority) as they are reported in the published studies. The trials appear to be at a low risk of bias, according to assessment by relevant criteria for superiority and non-inferiority trials.

The evidence suggests that RTX at 4x375mg/m^2 is non-inferior to oral CYC in terms of induction of remission in adults with AAV and *de novo* disease, and superior to oral CYC in terms of remission in adults with generalised, "severe" AAV who have relapsed one time on CYC (RAVE). It also has a similar safety profile to oral CYC. The evidence also suggests that RTX at 4x375mg/m^2 plus 2-3 IV pulses of CYC is non-inferior to IV pulse CYC in terms of remission in adults with generalised, "severe" AAV and *de novo* disease (RITUXVAS). RTX plus IV pulse CYC also has a similar safety profile to IV pulse CYC alone. The evidence only relates to induction of remission with these specific regimens in adult populations with generalised, "severe" AAV and only the 4x375mg/m^2 dose is considered to be licensed and relevant. No evidence is presented on the efficacy or safety of RTX in adults with mild AAV, in children, or using this regimen as a maintenance therapy or for relapse after RTX, as these indications and therapies are all outside the licence and are therefore *"not something* [the manufacturer] *could legally promote"* (see manufacturer's Clarification Response: Priority request 6[20]). In addition, the RITUXVAS trial had few, and the RAVE trial no patients with the most severe forms of the disease.

The principal uncertainties relate to the limitations of the submitted evidence; these are discussed below.

Population:

The populations in the two principal trials are different. The RAVE trial only considers relatively young adults (mean 54 or 51 years for RTX and CYC arms respectively) with moderately severe AAV and either *de novo* disease or following relapse after CYC. It does not include adults with severe renal impairment, life threatening pulmonary haemorrhage, those contraindicated for CYC or those CYC-refractory. It is uncertain if RTX alone will demonstrate equal efficacy and safety in other adult populations or children. The RAVE trial's younger population potentially confounds some outcomes as age is a known predictor of relapse among AAV populations.[37] It is perhaps worthy of note that rates of complete remission with the prednisone taper (RAVE) at 6 months were more comparable between RTX and oral CYC for populations with renal involvement.[33] The RITUXVAS trial

considers a much older population (median 68 or 67 years across arms), with severe renal impairment, and including those intolerant to CYC.

Intervention:

There are two possible dosing regimens for induction of remission in this population: 4x375mg/m^2 (the licensed dose) and 2x1g (the commonly used dose, see MS, Table 8). The effectiveness evidence submitted only concerns the 4x375mg/m^2 dose for induction therapy for *de novo* or CYC-relapsed generalised, "severe" AAV patients. It is uncertain whether the 2x1g dose, currently used for RA[25] and evaluated in non-RCT studies[38,41,49] would demonstrate equivalent or superior efficacy and/or safety for this population. The 2x1g dose is described variously in the MS as the "commonly used dose", the "incorrect dose" and the "off-label" dose and appears to show equivalence[39,49] or superiority[41] in retrospective studies of efficacy and safety when compared to the 4x375mg/m^2 dose in AAV patients with refractory of relapsing AAV. The 2x1g dose is actually a smaller overall dose than the 4x375mg/m^2 dose[25] which might explain its better safety profile. The relative efficacy and safety of the 2 regimens has not been assessed in a RCT, so there is clinical equipoise in terms of the optimal dosing regimen for RTX as an induction therapy in this population.[4,25] It is not clear why licensing was sought only for the 4x375mg/m^2 dose, particularly as, according to clinical advice, CD 19 counts are ablated by both regimes in the majority of patients for > 6 months. There is therefore uncertainty concerning the appropriateness of the 4x375mg/m^2 dose for this AAV population. This is especially important for two reasons. Firstly, the rationale for alternative therapies, such as RTX, is the need to reduce exposure to the various therapies in this population.[4,23-25] Secondly, it is recognised in the MS (see Section 2.6) that the lower 2x1g regimen is currently the most widely-used off-label dose in the UK for severe AAV.

The question whether RTX or RTX+CYC offers the optimum intervention is also uncertain and has been raised in the literature:[4,25] the remission rates and safety profile achieved in older, more severe populations in the RITUXVAS trial and the Mansfield *et al.*[38] study suggest that this combination therapy might have a role to play in induction of remission in patients with *de novo* AAV. In the absence of a head-to-head trial it is therefore also uncertain whether RTX (at either 4x375mg/m^2 or 2x1g) + CYC is inferior, equivalent or superior to RTX alone at the dose of 4x375mg/m^2 in terms of safety and efficacy.

Comparators:

The submission focuses on RTX against oral CYC as a comparator. It is uncertain whether RTX without CYC would demonstrate equal efficacy and safety if it were compared with IV CYC, which might have a better safety profile[41,49] and which clinical advice suggests is used more often in expert clinical practice. It is also uncertain how RTX compares with other potentially relevant comparators

specified in the scope and decision problem, such as MMF and MTX, which might all be used for similar populations (adult patients with generalised, "severe" AAV with and without severe renal impairment respectively). Some evidence is available which could have informed an indirect comparison, however the submission did not contain such an analysis.

Outcomes:

There are uncertainties concerning efficacy and safety in the specified population beyond 6-18 months,[4,25] the duration of the longest trial. Duration of remission is not reported as an outcome in RAVE or RITUXVAS and the median time to relapse with RTX and CYC might be more than 12 months.[25] The RAVE and RITUXVAS trials report non-inferiority to CYC for complete remission (different definitions for remission and different regimens for CYC), and similar adverse event profiles, at 6- and 12-months respectively. However, reduced adverse events might have been expected and was part of the hypothesis for the RITUXVAS trial. This should also be the case for the RAVE trial which employed oral CYC as the comparator, a regimen with a possibly worse adverse event profile than the IV pulse CYC regimen employed in the RITUXVAS trial. The results were most relevant for the number of malignancies in the RTX arm of the RAVE trial and the number of deaths in the RTX arm in the RITUXVAS trial, although these may well have been pre-existing and/or causative of the vasculitis, given the relatively short follow-up period; while concerns about the number of deaths in the RTX arm in the RITUXVAS trial were examined and addressed.

One reason suggested for the comparable rather than superior safety profile in the RITUXVAS and RAVE trials relates to their short study durations (therefore fertility effects might not been seen[23]). However, the RAVE trial compared RTX with the "less safe" oral regimen of CYC and used a relatively young population with less severe symptoms. A further reason might be the high cumulative dose with RTX in both trials,[4,5,38] which might be causing adverse events. Finally, it is possible that the 4x375mg/m^2 dose of RTX is contributing to higher than expected rates of adverse events in both trials because this is a higher overall dose than the 2x1g dose.[4,25] The 2x1g dose was used in a long-term study by Mansfield *et al.*[38] which reported few adverse events, despite involving an older population with severe symptoms.[38] The hypothesised short-term benefits of fewer adverse events for RTX has therefore not been demonstrated by the evidence submitted.

RTX regimens did not demonstrate superiority to CYC in terms of adverse events, which was a hypothesis of at least the RITUXVAS trial, and several adverse events were more frequent in the RTX arms (deaths and malignancies, albeit these were not significant differences and the relationship with malignancy in relatively short-term studies is complex, as above). Longer duration trials with more comparable groups are needed to assess malignancies and fertility outcomes.

Study design and quantity:

There are uncertainties concerning efficacy and safety in the specified population because the evidence consists only of two short-duration RCTs for slightly different populations (mean ages of 54 and 51 years in the two RAVE arms, relatively moderate renal involvement and CYC tolerant only [RAVE] vs mean age 68 years, AAV with more severe renal involvement and including potentially CYC intolerant populations [RITUXVAS]) using different interventions (RTX or RTX+CYC) and comparators (by dose and administration). The evidence base is therefore very limited in size.

Quality of evidence:

RITUXVAS is generally at low risk of bias across most domains, but is at high risk of performance bias and detection bias due to being open-label in design and due to lack of clarity on methods used for blinded outcome assessment. This presents a potential confounder as the primary endpoint is measured using a scale that might possibly involve a degree of subjectivity (BVAS/WG),[42,43] but otherwise has shown good inter-observer agreement for GPA populations.[44]

The RAVE trial appears to be at low risk of bias across all domains, however the ERG have a number of concerns regarding its design. Two protocols are available, one of which only exists in a version that post-dates (2009) the inception of the trial. This latter protocol is the only source of information on randomisation and allocation concealment; the source of the criterion of non-inferiority for the primary efficacy endpoint is only explained in the MS. The primary endpoint was updated during the course of the trial and the source of the prednisone dose threshold (<10mg/d) for one designated secondary efficacy endpoint is unclear. Subgroup analyses are described as "pre-specified" but details are not provided in any of the available protocols. Both trials might therefore be considered as being at low risk of bias, but such an assessment must take into account these caveats.

Summary

The MS only presents evidence on:

- An indication heavily limited by the licence (not including RTX as a maintenance treatment or for relapse other than post-CYC);
- A RTX dose that is not currently the common off-label dose in the UK;
- A single trial offering evidence on an alternative dose or regimen (RITUXVAS, RTX+CYC)
- A single trial using the dose and regimen that is to be licensed (RAVE);
- Data for only 6-12 months in the included trials, i.e. longer-term efficacy and safety outcomes are unknown;
- Some potential questions concerning certain adverse events, especially rates of mortality and malignancies.

5. COST EFFECTIVENESS

This chapter presents a review of the available evidence relating to the cost-effectiveness of RTX in the treatment of ANCA-associated vasculitis. Section 5.1 presents a critique of the systematic review of existing economic analyses of RTX undertaken by the manufacturer. Section 5.2 provides a detailed description of the health economic model submitted by the manufacturer and the economic analysis presented within the MS.[8] Section 5.3 presents a critical appraisal of the manufacturer's model and associated economic analysis. Section 5.4 presents additional work undertaken by the ERG to examine the impact of the key issues identified through the critical appraisal. Section 5.5 presents a discussion of the key issues and uncertainties relating to the cost-effectiveness of RTX for the treatment of ANCA-associated vasculitis in England and Wales.

5.1 ERG comment on manufacturer's review of cost-effectiveness evidence

The manufacturer conducted a systematic review with the objective of identifying studies that addressed the cost-effectiveness of one or more interventions for patients with either GPA or MPA. The manufacturer undertook systematic searches across the following electronic databases and research registers:

- Medline
- Medline In-Process
- Embase
- Embase Alerts
- Econlit
- NHS EED

Articles were excluded if they were not related to humans, not written in the English language, or the patient population was inappropriate (≤18 years old) (see MS,[8] p.112). Articles were also excluded if they reported on the cost-effectiveness of treating co-morbidities potentially associated with MPA or GPA. There was no restriction in the search strategy with respect to intervention and comparators.

Articles were included if they reported a measure relevant to cost-effectiveness, for example:

- incremental cost per QALY gained
- cost of being in remission/not being in remission
- cost of disease relapse compared with not having a relapse
- direct and indirect costs of treating GPA or MPA in any currency and at any geographical location.

Articles were also included if they used a decision model to estimate the cost-effectiveness of any intervention for MPA or GPA. A total of 159 records were identified, however ultimately all were excluded. Thus, the systematic review did not identify any studies that reported on the cost-effectiveness of treatment for MPA or GPA.

In addition, the manufacturer conducted systematic reviews in an attempt to identify studies investigating HRQoL and resource use in patients with GPA or MPA, but reported that no relevant studies were found.

The manufacturer's reporting of search strategies for finding cost-effectiveness, HRQoL and resource use evidence was adequate. The minimum required sources were searched. Study design filters were applied by the manufacturer. As described in the clinical effectiveness review critique of the manufacturer's searches (Section 4.1.1), the sensitivity of the economic evaluation search strategies could have been improved by the inclusion of free-text synonyms for "rtx" in all search strategies i.e. "rituximab" or "mabthera", "rituxan" and "rituxin" and the subject heading "Vasculitis". The ERG considered that the language restriction to English only publications was too restrictive and it is not clear whether relevant foreign language publications have been missed. Translation of the search strategies from Medline and Embase to other databases was not consistently applied; intervention terms were omitted in the NHS EED and EconLit searches.

The ERG is not certain that all potentially useful evidence was identified and appropriately incorporated in the economic evaluation by the manufacturer. For instance, for the main cost-effectiveness search, the manufacturer stated that studies would be included even if they were restricted to providing evidence on the costs associated with treating MPA or GPA, and no-matter which country or geographical location they provided evidence for. However, the manufacturer goes on to report that two studies were deemed worthy of further inspection but were subsequently excluded for not being cost-effectiveness analyses (see MS,[8] p.115). These two studies (Hoffman *et al.* (1998[50]) and Ndir *et al.* (2011[51])) were US and French costing studies that investigated the direct and indirect costs associated with GPA.[50,51] While these may not have provided evidence relevant for a UK perspective this is not clear from the MS, and it is not clear why these were excluded from the manufacturer's review. Given the importance of health state cost estimates used in the manufacturer's economic model – which were based purely on assumption informed by the opinion of one clinical advisor to the manufacturer – this is particularly relevant.

5.2 Description of the manufacturer's economic model

5.2.1 Model scope

The model presented by the manufacturer estimates the incremental costs and health effects of a treatment sequence beginning with RTX compared to a treatment sequence that begins with CYC. The sequence that begins with CYC is assumed to represent the "standard of care" whereas the RTX sequence represents a pathway of care deemed by the manufacturer to be realistic based upon expert opinion (see MS p.119). Cost-effectiveness is presented in terms of the incremental cost per quality-adjusted life year (QALY) gained from the perspective of the NHS and Personal Social Services (PSS) over a lifetime time horizon. In practice, only NHS costs are included. Three populations can be evaluated within the model: (1) treatment naïve (2) recurrent disease and (3) all patients. The "all patients" population (which forms the manufacturer's base case analysis) is made up of "treatment naïve" and "recurrent disease" patients, however the structure of the model and the parameter values used for the different populations means that the "all patients" analysis does not represent an average of the "treatment naïve" and "recurrent disease" populations – this will be discussed in Section 5.3. The manufacturer's model only considers the use of RTX as an induction therapy for these populations – maintenance therapy or the use of RTX following an initial relapse observed in the model is not considered. While the model allows a subgroup analysis of a "recurrent disease" population – that is, patients who have previously been treated but have relapsed – patients who relapse within the model are not permitted to receive RTX, even if they enter the model as "treatment naïve" patients.

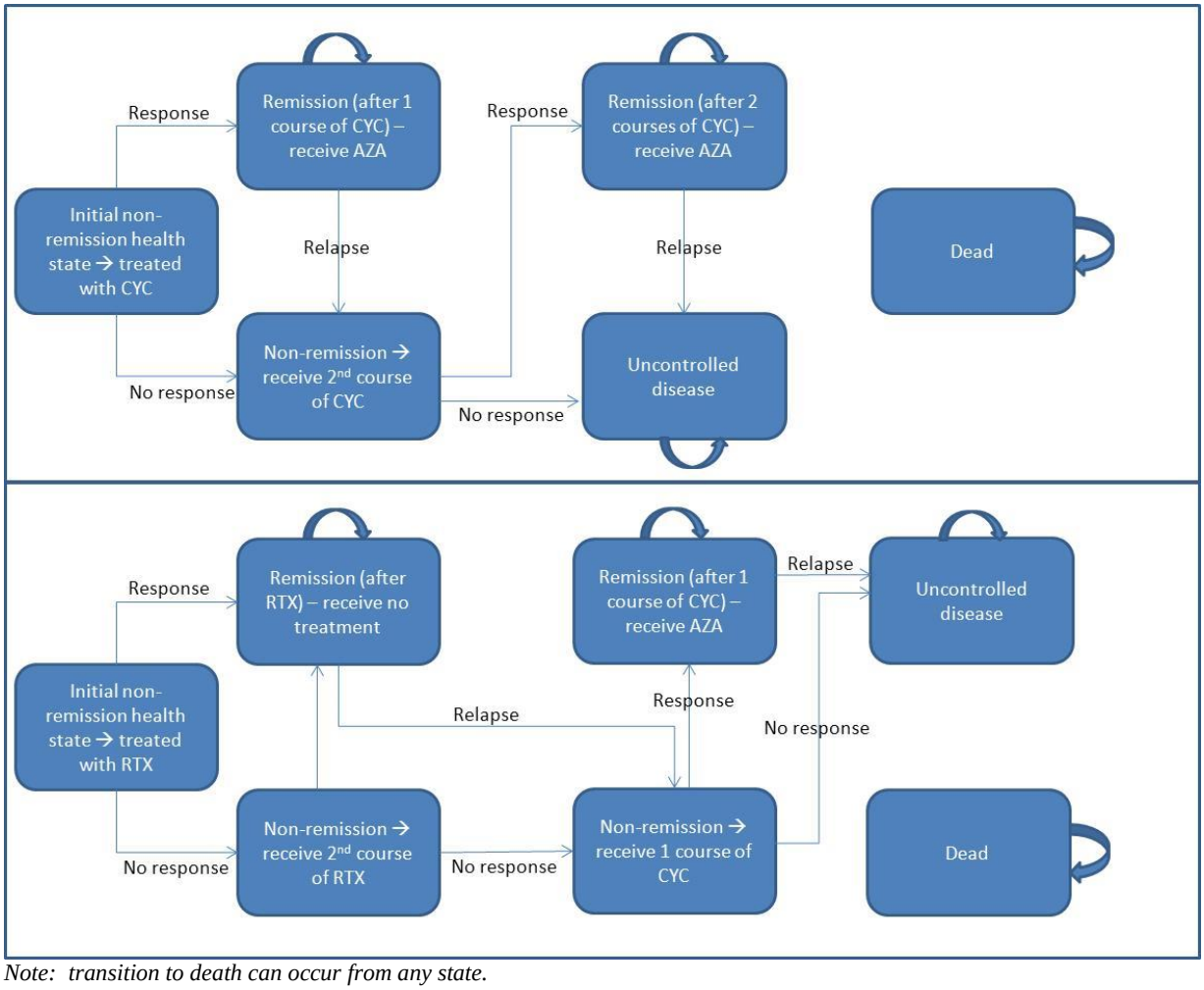
In line with current methodological guidance,[52] all costs and health outcomes are discounted at an annual rate of 3.5%. The model was programmed in Microsoft Excel® with an additional macro written using Visual Basic for Applications (VBA) to perform probabilistic sensitivity analysis (PSA).

The manufacturer's submission[8] lacked some clarity regarding the disease and treatment pathway assumed within the economic model, and regarding the values assumed for some key parameters. In order to ensure clarity regarding the manufacturer's modelling methods, assumptions and the data that underpin the model, Sections 5.2.2, 5.2.3 and 5.2.4 provide a detailed description of the submitted model. This description has been produced by the ERG through a detailed scrutiny of the submitted model, the MS report and subsequent clarification response.

5.2.2 Model structure

A conceptual form of the model implemented by the manufacturer is presented in Figure 1, as produced by the ERG. This illustrates the disease pathway and the associated treatment sequence for the RTX and CYC groups.

Figure 1: Rituximab model structure (drawn by the ERG)



Note: transition to death can occur from any state.

Although the model takes a Markov structure, multiple "remission" and "non-remission" health states are included in order that the desired treatment pathway can be modelled. The treatment received at any point in the disease pathway depends upon the number of courses of treatment previously received. "Non-remission" states are tunnel states, in which patients only remain for one cycle of the model, before either responding or not to the treatment they receive. The costs (other than treatment costs) and utilities associated with each "remission" state are the same, but separate states are modelled such that a specific treatment sequence can be modelled for patients initially treated with CYC and patients initially treated with RTX. Similarly, the costs (other than treatment costs) and utilities associated with each "non-remission" health state are identical, but multiple states exist in order to allow treatment sequences to be modelled. The structure of the model and the treatment sequences received remain the same for each population modelled – the "treatment naïve", "recurrent disease" and "all patient" populations. However, the model has the capacity to alter the number of courses of CYC received in both treatment groups. It also has the capacity to alter the proportion of patients in the RTX group who receive a second course of RTX following initial failure.

The model generates a probability of residing in each health state at six month intervals. In the base case analysis, the probability of being in the death state reaches 1.00 after 78 cycles (39 years). The age at entry into the model is assumed to be 52.8 years, based upon the mean age at baseline observed in the RAVE trial.

Model structure for patients initially treated with cyclophosphamide

For patients initially treated with CYC only two courses of CYC are permitted in the base case (although the model has the capacity to test between 0 and 5 courses). It is assumed that 72% of patients receive IV CYC pulses (15mg/kg every 2 weeks for the first 3 pulses, followed by infusions every 3 weeks for the next 7 pulses) with the remainder receiving oral CYC (2mg/kg per day), based upon the manufacturer's market research (see MS p.161). Patients are initially treated with CYC (beginning in the "Initial non-remission health state → treated with CYC" health state in Figure 1) and either respond, do not respond, or die. Those patients who respond move into the "remission" health state and receive AZA (2mg/kg per day) while they remain in this state. Each cycle, patients can either remain in this state, die, or relapse back into non-remission. Patients who relapse move into a second "non-remission" health state and receive a second course of CYC – again from this point, patients can either respond, not respond, or die. Those patients who do not respond to this second course transit into the "uncontrolled disease" health state, whereas patients who do respond move into a second "remission" health state, in which AZA is again administered as maintenance treatment. Patients remain in this state until they die or relapse. When patients relapse from this second remission state they move into the "uncontrolled disease" health state because it is assumed that no more CYC treatment is available to them. Patients remain in the "uncontrolled disease" state until they die. Patients who are initially treated with CYC and do not respond move directly into the second "non-remission" health state and receive a second course of CYC – joining the pathway described above.

Model structure for patients initially treated with rituximab

For patients initially treated with RTX only one subsequent course of CYC is permitted in the base case (although the model has the capacity to test between 0 and 5 courses). Again it is assumed that when CYC is administered, 72% of patients receive IV CYC pulses (15mg/kg every 2 weeks for the first 3 pulses, followed by infusions every 3 weeks for 7 pulses (treatment continues for 6 months) with the remainder receiving oral CYC (2mg/kg per day for 6 months). The modelled RTX dosing schedule is one weekly dose of 375mg/m^2 for 4 weeks. The manufacturer's base case assumes that two courses of RTX are administered to patients who do not respond to initial treatment, but that otherwise only one course is given. The model has the capacity to alter this assumption such that three other scenarios can be modelled: (i) no patients receive two courses of RTX treatment; (ii) only patients who respond and then relapse receive two courses or; (iii) all patients receive two courses.

Patients are initially treated with RTX (beginning in the "Initial non-remission health state → treated with RTX" health state in Figure 1) and either respond, do not respond, or die. Those that respond move into the first "remission" state. While in this state they do not receive any maintenance treatment, in line with the treatment strategy tested in the RAVE trial. Each cycle, patients can either remain in this state, die, or relapse back into non-remission. Patients who relapse move into what we describe here as a "third" "non-remission" health state and receive a course of CYC – again from this point patients can either respond, not respond, or die. Those that do not respond transit into the "uncontrolled disease" health state, whereas those that respond move into a second "remission" health state – in this state AZA is administered as maintenance treatment. Patients remain in this state until they die or relapse. When patients relapse from this second remission state they move into the "uncontrolled disease" health state because it is assumed that no more CYC treatment is available to them. Patients remain in the "uncontrolled disease" state until they die. Patients who are initially treated with RTX and do not respond move directly into a second "non-remission" health state and receive a second course of RTX. Again from this point they either respond, do not respond or die. Those that respond move into the first "remission" health state – joining the pathway described above. Those that do not respond move into the third "non-remission" health state and receive a course of CYC, joining the pathway described above.

The treatment sequences incorporated within the manufacturer's model essentially mean that RTX is modelled as an additional therapy within the treatment pathway for a proportion of the population. Whereas patients initially treated with CYC can only receive two courses of treatment before entering the "uncontrolled disease" health state, patients who initially do not respond to RTX receive a second course of RTX and a subsequent course of CYC – hence this group of patients receive three lines of treatment before entering the "uncontrolled disease" state. The implication of this is that the earliest that patients in the RTX group can enter the "uncontrolled disease" health state is the third cycle of the model (following non-response to initial RTX treatment (cycle 1), non-response to a second course of RTX (cycle 2), and non-response to one course of CYC (cycle 3)). In the CYC group, the earliest point at which patients can enter the "uncontrolled disease" state is cycle 2 (following non-response to an initial course of CYC (cycle 1) and non-response to a second course of CYC (cycle 2)). Hence RTX adds a line of therapy for non-responders, delaying the time to uncontrolled disease for these patients. Patients in the "uncontrolled disease" health state are assumed to have an increased risk of death and a lower utility score than patients in the "non-remission" health states (in each state mortality risks and utility scores are age-dependent), and patients in the "non-remission" health states are assumed to have an increased risk of death and a lower utility score than patients in the "remission" health states. Hence increasing time spent in "remission" health states and reducing time spent in "uncontrolled disease" affords health benefits.

Mortality risks within the model are driven by age and sex based upon UK life tables,[53] adjusted according to the mean age and proportion of males and females recruited in the RAVE trial and a published standardised mortality ratio (SMR) comparing a general population to an AAV population (see MS p.132).[54] Assumptions are made regarding the reduced and increased mortality risks associated with the "remission" and "uncontrolled disease" states respectively, with the AAV SMR applied directly to the "non-remission" health states.

HRQoL is based upon age and sex using published UK-based evidence for the general population,[55] adjusted according to the RAVE demographics and weighted for "remission" and "non-remission" health states according to SF-36 data (subsequently transformed into EQ-5D utility scores) collected from the RAVE trial. An additional utility decrement is applied to patients in the "uncontrolled disease" health state, and further decrements are made according to the probability of experiencing adverse events in each health state. Total QALYs are calculated as a function of the time spent in each health state and the associated utility score.

The model includes drug acquisition and administration costs, monitoring costs, costs of outpatient appointments, and costs of managing adverse events.

The treatment received while in the "non-remission" health states influences the probability of response, the probability of relapse if a response is achieved, and the proportion of patients that are assumed to experience an adverse event. Increased response rates and reduced relapse rates lead to QALY gains and reductions in future costs (because the "remission" health state is assumed to be much less costly than "non-remission" and "uncontrolled disease" health states). RTX and CYC are assumed to have different response rates associated with both first and subsequent courses. Relapse rates differ for remissions induced by the two treatments but in the base case model the same relapse rate is applied for first and second relapses (though in the analyses of the "treatment naïve" and "recurrent disease" sub-populations different relapse rates are applied). The proportion of patients that experience adverse events also differs according to treatment received. Hence response rates, relapse rates, adverse event rates and the additional step in the treatment pathway incorporated within the RTX treatment sequence determine the incremental health gain attributed to RTX. These factors, combined with the different drug acquisition and administration costs, monitoring costs and management costs, drive the incremental cost increase attributed to RTX. Using these incremental costs and QALYs an incremental cost effectiveness ratio is calculated.

5.2.3 Key assumptions underpinning the model

The manufacturer's base case model makes the following key structural assumptions:

- RTX is given using a dosing regimen of 375mg/m^2 once a week for four consecutive weeks.

- Patients who do not respond to initial RTX treatment are immediately retreated with a second course of RTX.
- Patients who achieve remission induced by RXT receive no maintenance therapy with any treatment.
- Patients who relapse after a RTX-induced remission, or who do not respond to an initial two courses of RTX, receive one course of CYC treatment.
- RTX is only considered as an initial induction treatment for each population ("treatment naïve", "recurrent disease" and "all patients"). It is not considered as a secondary induction treatment after a relapse observed in the model.
- Response rates, relapse rates and adverse event rates are identical for CYC IV pulse treatment and CYC oral treatment.
- The number of CYC courses received is not dependent upon the method of administration (i.e. IV pulse or oral).
- Patients are assumed to receive "full" treatment courses – that is, 4x375mg/m^2 for RTX, 2mg/kg/day for 6 months for oral CYC, and 10 administrations of 15mg/kg for IV CYC.
- Patients who initially receive CYC treatment may only receive two courses of CYC (this is assumed to be the case for all populations: "treatment naïve", "recurrent disease" and "all patients").
- Patients who achieve remission induced by CYC receive AZA maintenance therapy.
- Remission is defined by a BVAS/WG score of 0 and a successful taper of prednisone at 6 months (zero treatment with prednisone at 6 months).
- "Partial remission", for instance where the BVAS/WG score is 0 but prednisone treatment has not been completely stopped at 6 months, is not modelled and therefore is assumed to be similar to non-remission and the associated treatment is given.
- All relapses are treated as being identical – severe *and* limited relapses observed in the RAVE trial are combined to estimate the relapse rates used in the manufacturer's model – it is therefore assumed that all relapses lead to re-induction treatment.
- Treatments that may induce remission are restricted to RTX and CYC. After exhausting the specified number of courses with these treatments patients remain in an uncontrolled disease state until death.
- It is assumed that patients in the "uncontrolled disease" health state have highly symptomatic disease and receive only best supportive care.

The manufacturer's base case model makes the following key assumptions around parameter values:

- Patients treated with a second course of RTX have a reduced probability of achieving remission equivalent to the relative reduction observed between treatment naïve and recurrent disease patients treated with CYC in the RAVE trial.
- Relapse rates are higher in the "recurrent disease" population compared to the "treatment naïve" population for both RTX and CYC, based upon the RAVE trial. However within these analyses and within the "all patients" analyses the relapse rates remain the same irrespective of whether patients are in first or second remission.
- Relapse rates are assumed to follow an exponential distribution and are hence modelled as being time-independent.
- Patient demographics (age and sex) are based upon the RAVE trial, except for BSA and weight. The average BSA is assumed to be $1.79m^2$ in the model,[56] based upon UK cancer patients, rather than the $2.00m^2$ observed in the RAVE trial.
- The SMR reported by Lane *et al.*(2005)[54] is assumed to represent the mortality risk associated with patients who are not in remission. An arbitrary ±10% has been applied to the SMR to estimate the risk of death in the uncontrolled and complete remission health states.
- The decrement in utility between the "non-remission" and "uncontrolled disease" health states is assumed to be identical to the decrement between the "remission" and "non-remission" health states.
- The only adverse events assumed to be relevant for incorporation in the model are anaemia, leukopenia, deep vein thrombosis, dyspnoea, diarrhoea and pneumonia. Each of these (apart from diarrhoea which is attributed no utility decrement) is attributed a utility decrement that is applied according to the proportion of patients who experienced these events for the duration of time spent in the relevant health states (it is implicitly assumed that these events are resolved after 6 months). That is, AEs associated with RTX are only applied when patients are in health states in which RTX is currently being received. AEs associated with CYC are only applied when patients are in health states in which CYC is currently being received. It is implicitly assumed that the utility decrement applied for each AE is equivalent to an average utility that a patient with that AE would experience over a 6 month period. Patients in remission after RTX treatment are assumed to experience no AEs because they receive no treatment, whereas patients in remission after CYC are assumed to experience adverse events associated with AZA treatment. Patients in the "uncontrolled disease" health state are assumed to experience no AEs.
- The risk of AEs remains the same no matter whether patients have been previously treated.
- Anaemia, leukopenia, deep vein thrombosis and pneumonia AEs are assumed to lead to costs incurred by the NHS. Dyspnoea and diarrhoea adverse events are assumed to lead to zero costs to the NHS.

- The drug costs incorporated within the economic model assume that the full RAVE protocol doses of RTX, CYC and AZA are given, rather than being based upon the doses actually received in the trial. However, for methylprednisolone and prednisone the mean dose received in both groups included in RAVE combined is used, and this is assumed to be applicable to both the RTX and CYC groups.
- The base case analysis assumes that there is no wastage of drugs.
- It is assumed that the cost of intravenously administering CYC pulse therapy is equivalent to the infusion cost of RTX. Given the source of the infusion cost, it is assumed that the cost of infusing RTX and CYC is equivalent to the cost of infusing infliximab as estimated in the year 2000 (though this cost is uplifted to 2012 prices (see MS p.155).
- It is assumed that there is no administration cost associated with methylprednisolone in either the RTX or CYC groups, because it is assumed that this will be given on the same day. The MS states that this is a conservative assumption because not all CYC treatment will be given intravenously, in which case administration of methylprednisolone would generate a separate cost. It is worthy of note that in the RAVE trial patients received 1-3 pulses of methylprednisolone prior to RTX or CYC treatment and so some patients may generate additional administration costs in both treatment groups.
- It is assumed that costs associated with monitoring patients during treatment with RTX and IV CYC are captured within the administration cost applied for these treatments. Monitoring costs are applied for patients taking oral CYC and for those taking AZA as maintenance therapy – these include 1.5 blood tests and 1.5 liver function tests per month.
- Patients taking oral or IV CYC are assumed to receive pneumocystis jiroveci prophylaxis based upon BSR guidelines (see MS p.152). This is assumed to consist of 400mg of trimethoprim. It is assumed that this prophylaxis is not given to patients treated with RTX (though this is recommended in the draft SmPC for RTX and this is stated on p.19 of MS).[8]
- It is assumed that patients attend regular outpatient appointments, the frequency of which are determined by the health state that they are in. The type of consultant seen is associated with the proportion of patients with different organ involvement at baseline in the RAVE trial. Rheumatologists, nephrologists, pulmonologists, neurologists, otolaryngologists, ophthalmologists and dermatologists are all included. For instance, 61% of patients are assumed to visit a rheumatologist, 66% are assumed to visit a nephrologist, and so on (see MS p.157). These proportions sum to more than 1.00, because patients often had more than one organ involved in their disease. It is assumed that patients in the "non-remission" state attend appointments once every 1.5 weeks, but because of multi-organ involvement it is essentially assumed that patients in this state have 3.1 outpatient appointments every 1.5 weeks (equivalent to 53.8 appointments in a 6-month period). It is assumed that patients in the

"remission" health state attend appointments once every 3 months, but because of multi-organ involvement it is essentially assumed that patients in this state have 3.1 outpatient appointments every 3 months.

- In addition it is assumed that patients with pulmonary symptoms (53% of patients) receive chest X-rays or CT scans once every 1.5 months while in the "non-remission" and "remission" health states.
- In the "uncontrolled disease" health state it is assumed that patients attend one "specialist palliative care" outpatient appointment (NHS reference cost SD04A) every 1.5 weeks (equivalent to 17.4 appointments every 6 months).

5.2.4 Evidence used to inform the model parameters

Population subgroup characteristics

The initial characteristics of the "all patient" model population (which forms the main analysis undertaken by the manufacturer), and the "treatment naïve" and "recurrent disease" populations which are included as sub-group analyses, are shown in Table 9.

Table 9: Initial patient characteristics for patient subgroups and the overall population

Variable	All patients	Treatment naïve	Recurrent disease	Source
Age	52.8	52.8	52.8	RAVE trial
Percent female	49.7%	49.7%	49.7%	
Weight (kg)	67.2	67.2	67.2	Unclear
Height (cm)	171.5	171.5	171.5	RAVE trial
Body surface area	1.79	1.79	1.79	Sacco *et al.* (2010)[56]

It is worthy of note that each population is assumed to display the same baseline demographics, no matter whether patients are treatment naïve or have recurrent disease. This is despite the fact that data on these subgroups would have been available from the RAVE trial. In addition, the ERG could not identify the source of the assumed weight – this is not from the RAVE trial, and the ERG were unable to identify a figure in the Sacco *et al.* (2010[56]) publication.

Summary of key model parameters

All key model parameter values are presented in Table 10.

Table 10: Model parameters and distributions (all subgroups unless indicated)

Parameter	Distribution	Mean	Standard error	Source and notes
Remission rates				
CYC (1st course)	Beta	All patients=0.5306 Treatment naive=0.6458 Recurrent disease=0.4200	Not stated	RAVE trial, except for the 0.3065 value for the 2nd course of CYC treatment in the "recurrent disease", which is referenced to Hoffman *et al.* 1992.[15] However the ERG cannot identify this figure from the Hoffman *et al.* paper.
CYC (2nd course)	Linked to remission rate relative risk	All patients=0.4200 Treatment naive=0.4200 Recurrent disease=0.3065	Not stated	
CYC 2nd course remission rate relative risk (relative risk comparing 2nd course remission rate to 1st course remission rate)	Beta	All patients=0.7915 Treatment naive=0.6503 Recurrent disease=0.7298	All patients=0.1583 Treatment naive=0.1301 Recurrent disease=0.146	For "all patients" and "treatment naïve" these are derived in order to reflect remission rates observed in RAVE. For "recurrent disease" Hoffman *et al.* 1992 is referenced,[15] but the source of this value remains unclear. Standard errors are assumed to be 20% of the mean.
RTX (1st course)	Beta	All patients=0.6364 Treatment naive=0.6042 Recurrent disease=0.6667	Not stated	RAVE trial.
RTX (2nd course)	Linked to remission rate relative risk	All patients=0.4138 Treatment naive=0.3929 Recurrent disease=0.4335	Not stated	Assumed, based upon remission rate relative risk.
RTX 2nd course remission rate relative risk (relative risk comparing 2nd course remission rate to 1st course remission rate)	Beta	All patients=0.6503 Treatment naive=0.6503 Recurrent disease=0.6503	All patients=0.1301 Treatment naive=0.1301 Recurrent disease=0.1301	Assumed, based upon relative risk between remission rates observed in "treatment naïve" patients and patients with "recurrent disease" treated with CYC in RAVE. Standard errors are assumed to be 20% of the mean.
Relapse rates				
CYC (1st course)	Beta	All patients=0.1496 Treatment naive=0.0704 Recurrent disease=0.1668	All patients=0.0360 Treatment naive=0.0369 Recurrent disease=0.0527	Based upon exponential models fitted to RAVE trial data. Standard errors are assumed to be 30% of the mean.*
CYC (2nd course)	Equal to relapse rates after 1st course	All patients=0.1496 Treatment naive=0.0704 Recurrent disease=0.1668	All patients=0.0360 Treatment naive=0.0369 Recurrent disease=0.0527	The model does not include parameters for relapse rates after subsequent treatment courses – implicitly these are assumed to remain constant.
RTX (1st course)	Beta	All patients=0.1647 Treatment naive=0.1199 Recurrent disease=0.3058	All patients=0.0373 Treatment naive=0.0469 Recurrent disease=0.0645	Based upon exponential models fitted to RAVE trial data. Standard errors are assumed to be 30% of the mean.*
RTX (2nd course)	Equal to	All patients=0.1647	All patients=0.0373	The model does not include parameters for relapse rates after

Parameter	Distribution	Mean	Standard error	Source and notes
	relapse rates after 1st course	Treatment naive=0.1199 Recurrent disease=0.3058	Treatment naive=0.0469 Recurrent disease=0.0645	subsequent treatment courses – implicitly these are assumed to remain constant.
Standardised mortality ratios by health state (equal for all treatments)				
Remission	Fixed	4.32	N/a	Assumed based upon Lane *et al.* 2005.[54] These are applied to age- and sex-based mortality rates from the general UK population.[53] (ONS, 2011)
Non-remission	Fixed	4.80	N/a	
Uncontrolled disease	Fixed	5.28	N/a	
Health state utilities (equal for all treatments)				
Remission	Beta	0.84	0.02	Post-hoc analysis of RAVE data, transforming SF-36 to EQ-5D. Note that utilities reduce each cycle (due to increasing age) according to Ara and Brazier's model.[55]
Non-remission	Beta	0.75	0.02	
Uncontrolled disease	Linked to difference between "remission" and "non-remission"	0.67	N/a	Assume that the difference between "non-remission" and "uncontrolled disease" is the same as the difference between "remission" and "non-remission".
Adverse event utilities				
Anaemia	Beta	0.63	0.19	Borg *et al.* 2008.[57] Standard error assumed to be 30% of the mean.
Leukopenia	Beta	0.70	0.17	Wang *et al.* 2008.[58] assumed to be equal to utility for neutropenia.
Deep Vein Thrombosis	Beta	0.69	0.04	Mathias *et al.*1999.[59]
Dyspnoea	Beta	0.82	Not stated	Van den Boom *et al.* 2001.[60]
Diarrhoea	Fixed	No decrement	N/a	Assume no decrement.
Pneumonia	Beta	0.21	Not stated	Sisk *et al.* 1997,[61] based upon Erickson *et al.* 1995.[62]
Adverse event probabilities (by treatment: CYC – CYC; AZA – azathioprine; RTX – RTX)				
Anaemia	Beta	CYC/AZA=0.035 AZA=0.000 RTX=0.035	Not stated	RAVE trial. Data for AZA from post-hoc analysis.
Leukopenia	Beta	CYC/AZA=0.035 AZA=0.030 RTX=0.023	Not stated	
Deep Vein Thrombosis	Beta	CYC/AZA=0.012 AZA=0.000 RTX=0.000	Not stated	
Dyspnoea	Beta	CYC/AZA=0.012	Not stated	

Parameter	Distribution	Mean	Standard error	Source and notes
		AZA=0.000 RTX=0.000		
Diarrhoea	Beta	CYC/AZA=0.000 AZA=0.000 RTX=0.023	Not stated	
Pneumonia	Beta	CYC/AZA=0.024 AZA=0.000 RTX=0.023	Not stated	
Drug costs				
RTX	Fixed	£1.75 per mg	N/a	BNF 2013.[63]
CYC (tablets)	Fixed	£0.004 per mg	N/a	BNF 2013,[63] 50mg, 100 tab pack.
CYC (IV)	Fixed	£0.011 per mg	N/a	BNF 2013,[63] average of 500mg and 1000mg price.
% CYC treatment given orally	Fixed	72.00%	N/a	Assumption – manufacturer's market research.
Azathioprine	Fixed	£0.002 per mg	N/a	BNF 2013,[63] 50mg, 56 tab pack.
Methylprednisolone	Fixed	£0.017 per mg	N/a	BNF 2013,[63] average of 1g and 2g vial price.
Prednisone	Fixed	£0.43 per mg	N/a	BNF 2013,[63] average price of 1mg, 30 tab pack, 2mg 30 tab pack, 2mg 100 tab pack, 5mg 30 tab pack, 5mg 100 tab pack.
Trimethoprim	Fixed	£0.0003 per mg	N/a	BNF 2013,[63] 200mg, 14 tab pack.
Administration costs				
RTX	Fixed	£180.29	N/a	Barton *et al.* 2004 (inflated to2012 prices).[64]
CYC (tablets)	Fixed	£0.00	N/a	No cost.
CYC (IV)	Fixed	£180.29	N/a	Barton *et al.* 2004 (inflated to 2012 prices).[64]
Azathioprine	Fixed	£0.00	N/a	No cost.
Methylprednisone	Fixed	£0.00	N/a	No cost.
Prednisone	Fixed	£0.00	N/a	No cost.
Trimethoprim	Fixed	£0.00	N/a	No cost.
Monitoring costs				
RTX	Fixed	£0.00	N/a	No cost.
CYC (tablets)	Fixed	£108.00	N/a	Assumption. 1.5 blood tests and 1.5 liver function tests per month.
CYC (IV)	Fixed	£108.00	N/a	Assumption. 1.5 blood tests and 1.5 liver function tests per month.
Azathioprine	Fixed	£108.00	N/a	Assumption. 1.5 blood tests and 1.5 liver function tests per month.
Methylprednisone	Fixed	£0.00	N/a	No cost.

Parameter	Distribution	Mean	Standard error	Source and notes
Prednisone	Fixed	£0.00	N/a	No cost.
Trimethoprim	Fixed	£0.00	N/a	No cost.
Proportion of patients visiting following consultants and receiving tests (same in each treatment group)				
Rheumatologist	Fixed	0.610	N/a	Based upon the proportion of patients recorded with the corresponding organ involvement at baseline in the RAVE trial. The proportion for chest X-ray or CT scans reflects the proportion with pulmonary involvement.
Nephrologist	Fixed	0.660	N/a	
Pulmonologist	Fixed	0.533	N/a	
Neurologist	Fixed	0.203	N/a	
Otolaryngologist	Fixed	0.584	N/a	
Ophthalmologist	Fixed	0.264	N/a	
Dermatologist	Fixed	0.183	N/a	
Other consultant	Fixed	0.061	N/a	
Chest X-ray or CT	Fixed	0.533	N/a	
Frequency of outpatient consultant visits by health state (same in each treatment group)				
Remission	Fixed	0.33 per month	N/a	Assumption – expert opinion.
Non-remission	Fixed	2.90 per month	N/a	Assumption – expert opinion.
Uncontrolled disease	Fixed	2.90 per month	N/a	Assumption – expert opinion. This reflects visits for specialist palliative care.
Frequency of chest X-ray and CT scans by health state (same in each treatment group) for those with pulmonary involvement				
Remission	Fixed	0.66 per month	N/a	Assumption – expert opinion.
Non-remission	Fixed	0.66 per month	N/a	Assumption – expert opinion.
Uncontrolled disease	Fixed	0.00 per month	N/a	Assumption – expert opinion.
Unit cost of consultant outpatient visits and tests				
Rheumatologist	Log normal	£125.68	0.15	NHS Trusts Reference Costs 2009-10:[65] Consultant led follow-up attendance non-admitted face to face, service code 410.
Nephrologist	Log normal	£150.53	0.12	Service code 361.[65]
Pulmonologist	Log normal	£131.12	0.15	Service code 340.[65]
Neurologist	Log normal	£139.61	0.17	Service code 400.[65]
Otolaryngologist	Log normal	£72.99	0.17	Service code 120.[65]
Opthalmologist	Log normal	£73.47	0.12	Service code 130.[65]
Dermatologist	Log normal	£86.37	0.16	Service code 330.
Other consultant	Linked	£125.68	0.15	Equal to “rheumatologist”.
Specialist palliative care	Log normal	£254.06	0.34	NHS Trusts Reference Costs 2009-10:[65] Specialist palliative care: Outpatient, currency code SD04A.
Chest X-ray or CT	Log normal	£29.08	Not stated	Barton *et al.* 2004 (inflated to 2012 prices).[64] Probabilistic value represents a multiplication of the “blood test” cost, based upon their relative deterministic means.

Parameter	Distribution	Mean	Standard error	Source and notes
Blood test	Log normal	£3.00	0.27	NHS Trusts Reference Costs 2009-10.[65] Haematology (excluding anti-coagulant services), pathology services currency code DAP823.
Liver function test	Log normal	£9.00	Not stated	Barton *et al.* 2004 (inflated to 2012 prices).[64] Probabilistic value represents a multiplication of the "blood test" cost, based upon their relative deterministic means.
Unit cost of adverse events				
Anaemia	Log normal	£363.66	0.22	NHS Trusts Reference Costs 2009-10.[65] Currency code SA03F.
Leukopenia	Log normal	£406.60	0.21	NHS Trusts Reference Costs 2009-10.[65] Currency code WA04U.
Deep Vein Thrombosis	Log normal	£387.34	0.19	NHS Trusts Reference Costs 2009-10.[65] Currency code QZ20Z.
Dyspnoea	Fixed	£0.00	N/a	Assumed no cost.
Diarrhoea	Fixed	£0.00	N/a	Assumed no cost.
Pneumonia	Log normal	£1,224.70	0.14	NHS Trusts Reference Costs 2009-10.[65] Currency code DZ11C.

**In a new version of the manufacturer's economic model submitted in response to clarification questions from the ERG the standard errors associated with these parameters were amended and calculated based upon the numbers of patients in the RTX and CYC groups in the RAVE trial and the mean relapse rate (using the method of moments technique).*

The following sections detail the evidence used to inform all parameters within the model.

Remission and relapse rates

Within the manufacturer's model, the principal driver of health benefits is associated with the time spent in the "remission" health state. There are two key factors within the economic model that determine the time spent in the "remission" state:

- The structure of the model, which determines the number of treatment sequences received in each treatment arm (and thus the number of times a patient may enter the "remission" state)
- Parameter values for treatment response rates, relapse rates, and mortality rates.

The structure of the manufacturer's model is shown in Figure 1. The treatment sequences incorporated within the model mean that RTX is modelled as an additional therapy within the treatment pathway for a proportion of the population. Whereas patients initially treated with CYC can only receive two courses of treatment, patients who initially do not respond to RTX receive a second course of RXT and a subsequent course of CYC – hence this group of patients receive three lines of treatment. The implication of this is that patients in the RTX group have one additional chance of entering the "remission" state, and their entry into the "uncontrolled disease" state is delayed by at least one model cycle. The manufacturer justifies the choice of modelled sequences based upon expert opinion (see MS p.118-119). CYC is known to be associated with important side effects that become increasingly likely with high lifetime cumulative doses. Based upon this, the manufacturer assumed that a maximum of two CYC doses would be administered. The manufacturer does not alter this assumption depending upon whether oral or IV CYC is given (these lead to different cumulative doses per course). The manufacturer also does not alter this assumption when different sub-populations are run through the model – "all patients", "treatment naïve" and "recurrent disease" groups all receive two courses of CYC in the respective base case analyses. The manufacturer notes that their assumption that patients who initially do not respond to RTX receive a second course was not tested in the RAVE trial, and is based upon expert opinion (see MS p.119). Hence the structure of the manufacturer's economic model – which itself is a key driver of the incremental health benefits associated with RTX – is seemingly based upon the opinion of one expert.

Parameter values for treatment response rates and relapse rates are mainly taken from the RAVE trial, although for some parts of the modelled treatment sequences data are not available and assumptions are made. Response rates for the first course of CYC or RTX treatment given are taken from Table 17 of the MS (see MS p.70) for the "all patients" population, and from Table 24 (see MS p.80) for the "treatment naïve" and "recurrent disease" subgroups.[8] These represent an intention to treat (ITT) analysis of the proportion of patients who achieved complete remission at Month 6, whereby any missing data was imputed using worst case imputation (WOCF). In this analysis data were imputed

for 3 patients in the CYC group and 1 patient in the RTX group who discontinued permanently from the study or withdrew from the study before month 6, without already having been classified as failing treatment. These patients were assumed to have failed treatment at 6 months. When these patients were instead excluded from the analysis the difference in response rates between the RTX group and the CYC group was slightly smaller (9.5% compared to 10.6%).

In Section 6.3.2 of the MS it is stated that patients who experienced severe disease flares or treatment failure between Visit V5 (Month 1) and Visit V8 (Month 6) were crossed over to the opposite treatment arm (see MS p.43).[8] According to Figure 5 in the MS this occurred in 5 (5.1%) patients in the RTX group and in 7 (7.1%) patients in the CYC group (see MS p.64).[8] The MS states that for the ITT analysis, efficacy endpoints were analysed according to the treatment assigned rather than the treatment received and therefore the remission rates at 6 months used in the economic model may be confounded by crossover (see MS p.59).[8] However, resulting bias appears likely to be small considering the proportion of switching was low and similar in the two treatment groups.

For the CYC group, remission rates associated with a second course of treatment are taken from the "relapsing disease at baseline" row of Table 24 of the MS (see MS p.80) for the "all patients" and "treatment naïve" analyses.[8] These data represent the complete remission rates at 6 months (ITT population with WOCF) for 50 patients who were recruited into the RAVE trial with relapsing disease at baseline. For the "recurrent disease" subgroup no data on remission rates associated with a second course of CYC can be derived from the RAVE trial, and so the manufacturer instead applied an additional decrement to the remission rate based upon a study by Hoffman *et al.* 1992, who followed a cohort of 158 US patients with Wegener granulomatosis vasculitis for periods varying between 6 months and 24 years.[15] However the ERG was unable to identify how the decrement was derived from the Hoffman *et al.*[50] paper.

The relapse rates used in the manufacturer's model were derived from the RAVE trial. Exponential models were fitted to data from patients who had experienced completed remission at 6-months in order to estimate the time-to-event for relapse (see MS p.136).[8] These models were used to derive 6-month probabilities for relapse. The use of exponential models is discussed further in Section 5.3. The MS only provides information on the fit of exponential models to the "all patients" data. Different relapse rates are applied for the "treatment naïve" and "recurrent disease" subgroups and the ERG assumes that these were estimated by applying exponential models to patients who achieved a complete response at 6-months who were recruited to the RAVE trial as treatment naïve patients and patients with recurrent disease respectively. In the manufacturer's economic model, the same relapse rate is applied after different courses of treatment within each model run – that is, in the "all patients" analysis a relapse rate of 0.1496 is applied after first and second courses of CYC, and a relapse rate of 0.1647 is applied after first and second courses of RTX. In the "naïve patients" analysis a lower

relapse rate is applied – this is 0.0704 for CYC and 0.1199 for RTX after first and second courses of treatment. In the "recurrent disease" analysis a higher relapse rate is applied – 0.1668 for CYC and 0.3058 for RTX after first and second courses of treatment. Hence previous relapse affects the relapse rates assumed across each of the subgroup analyses, but does not cause the relapse rate to change *within* the analyses (for example, in the "treatment naïve" subgroup analysis, patients who achieve remission, relapse, and then achieve remission for a second time, are allocated a lower relapse rate than patients achieving a first remission in the "recurrent disease" subgroup analysis). This issue is discussed further in Section 5.3.

It is noteworthy that for longer term efficacy analyses (beyond 6 months), the MS states that data from patients that switched treatments (5.1% in the RTX arm and 7.1% in the CYC arm) are treated as missing and WOCF was used to impute data (see MS p.61).[8] Such analyses could result in selection bias and this may have affected the estimation of relapse rates.

Mortality rates also affect how long patients are estimated to stay in each health state of the manufacturer's model. In the model, mortality rates do not differ depending upon which treatment is being taken, however these do differ by health state. Background age- and sex-based UK general population mortality rates are used within the model,[53] but these are multiplied by a different SMR in each health state. The "base" SMR (4.8) is taken from a cohort study that identified 99 patients with ANCA-associated vasculitis in a region in England between 1988 and 2000 and compared death rates to those in the general local population.[54] Thirty-one deaths were observed within this study. Lane *et al.* do not classify patients as to whether they were in remission or not. In the MS it is assumed that the "base" SMR is applicable to patients in the "non-remission" health state and an arbitrary ±10% is applied to this for patients in the "remission" and "uncontrolled disease" health states respectively. The MS states that the assumption that patients in complete remission have a reduced risk of mortality compared to those not in remission is based upon UK registry data, but no further reference to these data are given (see MS p.120-121),[8] and no further evidence was provided in the manufacturer's response to the ERG's clarification questions. In the manufacturer's model the reduced mortality rate in the "remission" health state and the increased mortality rate in the "uncontrolled disease" health state leads to marginal survival benefits for the RTX group as these patients spend a longer duration in the "remission" health state and a shorter duration in the "uncontrolled disease" health state due to the assumed treatment sequences and the remission and relapse rates modelled.

Health related quality of life

In the RAVE trial, SF-36 data were collected at baseline and at 6 months. Summary scores improved at 6-months compared to baseline in both treatment groups and there was not a significant difference in this change between groups. In the manufacturer's economic model, utility scores do not depend upon the treatment being received, but do differ by health state. SF-36 scores collected in the RAVE

trial were converted into EQ-5D scores for the "remission" and "non-remission" health states using Ara and Brazier's mapping models.[66] It is not clear exactly which data from RAVE were used to inform these estimates – the ERG assumes that utilities for "remission" were estimated based upon data from patients who achieved complete remission at 6-months, but it is unclear whether utilities for "non-remission" included only data from non-responders at 6-months, or also included data from all patients at baseline. The MS states that Model 2 from Ara and Brazier was used for the base case economic analysis, as this model is reported to be more accurate when predicting incremental differences between study arms and changes over time (see MS p.140). Model 1 from Ara and Brazier was used in sensitivity analysis. Although the two models produced quite different utility scores for each state (0.79 vs. 0.84 for "complete remission" and 0.70 vs. 0.75 for "non-remission") the relative difference between the two health states was similar. Data were not available from the RAVE trial on HRQoL for the "uncontrolled disease" health state. In the manufacturer's economic model the utility of this state was assumed to be worse than the "non-remission" state by the same absolute amount that "non-remission" is worse than "remission" (see MS p.146).[8]

To apply utility scores within the economic model, the manufacturer first estimated general population utility given the age and sex observed in RAVE based upon Ara and Brazier's model.[55] These utilities were estimated for each 6-month cycle (utility decreases with age). The utilities estimated for "remission", "non-remission" and "uncontrolled disease" were then calculated as weights compared to the general population score which allowed utility scores to be estimated for each health state, for each cycle of the model. For instance, for a patient aged 52.8 years (the baseline age in the model) the general population utility is estimated to be 0.85 (given a 50.3% proportion of males), using Ara and Brazier's model.[55] In RAVE, the utility associated with being in remission was 0.84, and the utility associated with not being in remission was 0.75, equivalent to 0.98 and 0.88 as proportions of the general population utility, respectively. Therefore, throughout the economic model time frame the utilities associated with the remission and non-remission health states were equivalent to 0.98*general population utility, and 0.88*general population utility, respectively.

Adverse event data sources

Adverse events (AEs) have relatively little impact in the manufacturer's model. The adverse events included are anaemia, leukopenia, deep vein thrombosis, dyspnoea, diarrhoea and pneumonia. Data relating to the probabilities of these events occurring while taking CYC, RTX and AZA are presented in Table 10. The MS states that these were the only AEs that were Grade 3 severity or above that occurred in more than 2% of patients in any treatment arm in the RAVE trial (see MS p.131).[8] Following treatment discontinuation (that is, after achieving remission in the RTX group, and in the "uncontrolled disease" health state) it is assumed that no AEs occur.

The impact of these AEs on HRQoL was estimated based upon a variety of sources from the literature. For anaemia, a Swedish study by Borg *et al.* 2008.[57] was used in which a utility score for moderate anaemia was measured using time-trade-off. For leukopenia, the utility score was assumed to be equal to that for neutropenia, allowing a study by Wang *et al.* 2008[58] to be used, in which a utility score for leukopenia was estimated based upon SF-36 data on neutropenia collected in a US study.[58,67] For DVT, a US study by Mathias *et al.*1999 which measured a utility score for severe DVT symptoms using the Health Utilities Index was used.[59] For dyspnoea, a Dutch study by Van den Boom *et al.* 2001 was used, which measured utility using the Maastricht Utility Measurement Questionnaire, an adapted Dutch translation of the Health Utilities Index.[60] For pneumonia, a utility score was derived based upon a paper by Sisk *et al.* 1997,[61] which in turn was based upon a US study by Erickson *et al.* 1995.[62] Erickson *et al.* present the Healthy People 2000 study, which attempted to estimate years of healthy life in the US based upon perceived health status and activity limitations reported by approximately 250,000 respondents to the National Health Interview Survey. Erickson *et al.* allocated utility values to each perceived health status and activity limitation combination using multiattribute utility scaling, which led them to allocate a score of 0.21 to patients in "fair" health who were limited in their activities of daily living. Sisk *et al.* assumed that this represents an accurate description of a patient with pneumococcal bacteremia. However, it is worthy of note that Sisk *et al.* assume that this level of utility is only experienced for a duration of 34 days, whereas in the manufacturer's model the utility for each AE is assumed to be appropriate for a 6-month period. It should also be noted that these studies used different elicitation methods which is not ideal.

To incorporate the AE utilities within the economic model, the manufacturer first calculated the general population utility score for each of the AE source studies, given the mean age and sex included within each of the studies. The AE utility score as a proportion of the general population score was then calculated. This weight was then applied to the appropriate health states in the model according to the proportion of patients estimated to experience each event.

The cost impact of managing each AE was based upon relevant costs taken from NHS Trusts Reference Costs 2009-10.[65] For anaemia, the cost was associated with a day case for haemolytic anaemia without complications (currency code SA03F). For leukopenia, the cost was assumed to be that of a short-stay (four days or less) non-elective inpatient admission for acute febrile illness without complications (currency code WA04U). For DVT, the cost was assumed to be that associated with a short-stay, non-elective inpatient admission for DVT (currency code QZ20Z). For pneumonia, the cost was assumed to be that associated with a long-stay non-elective inpatient admission for lobar, atypical or viral pneumonia without complications (currency code DZ11C). Dyspnoea and diarrhoea were assumed to result in zero costs.

Unit cost data sources

Acquisition costs of the drugs included in the manufacturer's economic model were taken from the British National Formulary No. 64 (published in September 2012).[63] Costs were included for RTX, methylprednisolone and prednisone for patients treated with RTX, and CYC, AZA, methylprednisolone, prednisone and trimethoprim (as prophylaxis treatment in order to prevent Pneumocystis jiroveci) for patients treated with CYC. The total drug acquisition cost per cycle was calculated by multiplying the drug unit cost by the total dose. It is not clear that the most appropriate unit costs have been used in all cases – this is further discussed in Section 5.3. The manufacturer assumed that 72% of patients treated with CYC received IV treatment, with 28% receiving oral treatment, based upon unreferenced market research (see MS p.161).[8]

For RTX, CYC (oral and IV), AZA and trimethoprim, costs were based upon recommended or licensed doses rather than the doses observed in the RAVE trial (see MS p.154-155),[8] whereas for methylprednisolone and prednisone costs were based upon the average received in the RAVE study, with treatment groups combined (that is, it was assumed that the dose received was equal in patients treated with RTX and patients treated with CYC).

Administration costs were included for RTX and IV CYC. The costs of administration were based upon the cost estimated by Barton *et al.*,[64] for an administration of infliximab. The original reference presented a cost for a price year of 2000, and the manufacturer uplifted this to 2012 using the Hospital and Community Health Services Pay and Prices Index reported by Curtis (2012).[68] Barton *et al.* do not provide a source or a method for their administration cost estimate. The manufacturer assumes that the administration costs for RTX and IV CYC were equal, and implicitly that these were equal to the cost of administering infliximab. The manufacturer applied zero cost to the administration of methylprednisolone, assuming that this is administered at the same time as RTX (or CYC). This is assumed to add nothing to the cost of the administration, despite the assumption that only 72% of CYC is given intravenously.

Monitoring costs were included in the manufacturer's model for oral CYC and AZA. No additional monitoring costs were included for RTX or IV CYC because the manufacturer assumes that monitoring is undertaken during the administration of these drugs.

The manufacturer states that during oral CYC treatment patients are assumed to receive 1-2 blood tests and 1-2 liver function tests (LFTs) each month, based upon Lapraik *et al.* (2007).[9] In the manufacturer's model this is incorporated as an assumption that patients receive 9 blood tests and 9 LFTs in each 6 month cycle that treatment is received. The manufacturer states that during AZA treatment patients are assumed to receive 1-2 blood tests and 1-2 LFTs each month, based upon Chakravarty *et al.* (2008 and 2009).[69,70] In the manufacturer's model this is incorporated as an

assumption that patients receive 9 blood tests and 9 LFTs in each 6 month cycle that treatment is received.

The manufacturer assumes that the cost of a full blood test is £3, based upon NHS Reference Costs (currency code DAP823),[65] and that the cost of LFTs is £9, based upon uprating to current prices the cost of £6.19 reported by Barton *et al.* (2004).[64] Barton *et al.* estimated their cost based upon data provided by the University Hospital Birmingham NHS Trust in 2000. NHS Reference Costs report a cost of £1.29 for biochemistry pathology services (currency code DAP841[65]) which may be appropriate for LFT and renal function tests.

Separate from treatment administration and monitoring costs, the manufacturer assumes that patients attend regular outpatient appointments with relevant consultants, the frequency of which are related to the health state that they are in. The type of consultant seen is associated with the proportion of patients with different organ involvement at baseline in the RAVE trial (see MS p.157-158)[8], as described in Section 5.2.3. Rheumatologists, nephrologists, pulmonologists, neurologists, otolaryngologists, ophthalmologists and dermatologists are all included and are costed based upon the relevant NHS Reference Costs (Consultant led follow-up attendance, non-admitted face to face appointment).[65]

In addition, it is assumed that patients with pulmonary symptoms (53% of patients) receive chest X-rays or CT scans once every 1.5 months while in the "non-remission" and "remission" health states. The cost for these was assumed to be £29.08, based upon Barton *et al.* (2004) and uprated to 2012 prices. Barton *et al.* estimated their cost based upon data provided by the University Hospital Birmingham NHS Trust in 2000.[64] According to NHS Reference Costs, the cost of diagnostic imaging in a consultant led follow-up attendance, non-admitted face-to-face attendance is £18.56 (service code 812), while the cost of a CT scan of one area with no contrast in the outpatient setting is £100.41 (currency code RA08Z).[65]

In the "uncontrolled disease" health state it is assumed that patients attend one "specialist palliative care" outpatient appointment (NHS Reference Cost SD04A[65]) every 1.5 weeks for the remainder of their life (17.4 appointments per 6 month cycle) (see MS p.158).[8] This is reported to be based upon clinical opinion.

5.2.5 Methods for handling uncertainty

The MS includes details of one-way sensitivity analysis, structural sensitivity analysis around treatment sequences, subgroup analyses and probabilistic sensitivity analysis (PSA).

Probabilistic sensitivity analysis

Within the PSA, the following parameters were treated as uncertain quantities:

- Remission rates
- Relapse rates
- Adverse event rates
- Adverse event costs
- Resource use costs (tests, scans, outpatient appointments)
- Health state utilities
- Adverse event disutilities

Uncertainty around remission rates, relapse rates, adverse event rates and utilities was characterised using beta distributions. Uncertainty surrounding costs was characterised using log normal distributions. The uncertainty surrounding all other model parameters (resource use rates [that is, the number of outpatient appointments and tests/scans in each health state], standardised mortality ratios, drug costs, administration costs and monitoring costs) was not considered within the analysis and these parameters were fixed at their point estimates. One thousand probabilistic samples were used and results were presented using cost-effectiveness planes and cost effectiveness acceptability curves (CEAC). It is noteworthy that the headline cost-effectiveness results presented in the MS relate to point estimates of parameters rather than the expectation of the mean.

Simple sensitivity analysis

One-way deterministic sensitivity analyses were conducted for key model parameters, as shown in Table 11. The results of these analyses were presented as point estimates of the incremental cost effectiveness ratio (ICER), rather than expectations of the mean. Deterministic sensitivity analyses were conducted for the “all patient” analyses, not the subgroup analyses.

Table 11: One-way sensitivity analyses considered

Model parameter	Base-case analyses	Sensitivity analyses
Remission rates	Different for each treatment: 0.5306 for CYC 0.6364 for RTX	Equal for CYC and RTX (0.6364)
Remission rates	Reduced for second induction regimens: 0.4200 for CYC 0.4138 for RTX	No reduction: 0.5306 for CYC 0.6364 for RTX
Mortality rates	Different for each health state: 4.32 for complete response 4.80 for non-response 5.28 for uncontrolled disease	Equal in all health states (4.8)
Mortality rates	Health states differ by 10%: 4.32 for complete response 4.80 for non-response 5.28 for uncontrolled disease	Health states differ by more: 3.36 for complete response 4.80 for non-response 6.24 for uncontrolled disease
Uncontrolled disease utility	Assumed be less than utility for non-response by same amount that non-response is less than complete remission (0.67)	Assume same utility as non-response, and a larger decrement: 0.75 0.58
Wastage	No wastage	No vial sharing
% of patients treated with IV CYC rather than oral	72%	100% 50% 33%
Patient weight	67kg	87kg
Health state resource use	Assumed resource use per 6 month cycle: 6.2 for complete remission 53.8 for non-response 17.4 for uncontrolled disease	Change number of visits in each health state by: +50% -50%
RTX dose	375mg/m^2 once weekly for 4 weeks	1000mg on day 1 and day 15

In addition, as part of their clarification response,[20] the manufacturer undertook sensitivity analysis around AE event rates assumed within the model. The manufacturer tested the impact on cost-effectiveness results of assuming that CYC had the same AE rates as RTX, and assuming that CYC has slightly lower AE rates. In their clarification response the manufacturer also tested the sensitivity of the model to basing prednisolone and methylprednisolone costs on actual doses received by treatment arm in the RAVE trial, rather than using an equal weighted dose in each arm.

Finally, in their clarification response,[20] the manufacturer tested the sensitivity of the model results to the best supportive care cost applied in the "uncontrolled disease" health state. Instead of applying a cost of £254 per appointment based upon NHS reference cost SD04A, a cost of £125.68 was used, based upon the NHS Reference Cost for a consultant led follow-up face-to-face appointment with a rheumatologist (service code 410).

Structural sensitivity analysis

The manufacturer acknowledges that the greatest source of structural uncertainty in their economic model is the assumed treatment pathway (see MS, p.162). The model assumes that due to the cumulative toxicity effects associated with CYC it is reasonable to apply a "cap" on the number of courses of CYC received – this is set at 2 in the base case. The manufacturer tested this structural assumption in the following scenario analyses:

- Allow only 1 course of CYC in the control group of the model, and no courses of CYC in the RTX group.
- Allow 2 courses of CYC in the control group, but no courses of CYC in the RTX group.
- Allow 2 courses of CYC in the control group and in the RTX group.
- Allow 1 course of CYC in the control group and in the RTX group.

In addition, the manufacturer assumed that patients who initially do not respond to RTX treatment immediately receive a second course of RTX. This assumption was tested by examining the following scenarios:

- No patients in the RTX arm of the model receive a second course of RTX.
- All patients in the RTX arm of the model receive a second course of RTX.
- Only patients who respond to RTX and subsequently relapse are eligible for a second course of RTX.

The results of these structural sensitivity analyses are presented as point estimates of the ICER, rather than expectations of the mean. The full range of these analyses was conducted only for the "all patient" analyses, not the subgroup analyses.

Subgroup analyses

As noted above, the manufacturer has reported results of analyses for "all patients", "treatment naïve" and "recurrent disease" population groups. The "all patient" analyses forms the manufacturer's base case analysis, whilst the other analyses are treated as subgroup analyses. Hence, the full range of deterministic and structural sensitivity analyses is applied to the "all patient" analyses and not to the "treatment naïve" or "recurrent disease" analyses. For the "treatment naïve" and "recurrent disease" analyses, remission rates and relapse rates are adjusted in accordance with the data on treatment naïve and recurrent disease patients obtained from the RAVE trial. The MS notes that because treatment naïve patients represented approximately half of the patients recruited to the RAVE trial the "treatment naïve" and "recurrent" disease remission rates and relapse rates are estimated using approximately half the amount of data as is used in the "all patients" analyses. This amounts to approximately 50 patients in each treatment arm.

Some structural sensitivity analyses were undertaken on the subgroup analyses. In the "base case" subgroup analyses it was assumed that patients in the control arm of the economic model received 2 courses of CYC, and patients in the RTX arm received 1 course of CYC, in line with the base case "all patients" analysis. The following alternative assumption was tested for the "treatment naïve" analysis:

- 2 courses of CYC given in the control arm and the RTX arm.

The following alternative assumption was tested for the "recurrent disease" analysis:

- 0 courses of CYC given in the control arm and the RTX arm.

In all of these analyses it was assumed that an additional course of RTX was given to patients who initially did not respond to RTX treatment.

5.2.5 Cost-effectiveness results reported within the MS

Central estimates of cost-effectiveness (as presented within the MS[8])

Table 12 presents the headline cost-effectiveness results for the modelled RTX sequence of treatments compared to the modelled CYC sequence of treatments. Results for the base case "all patients" analysis are presented, as well as results for the "treatment naïve" and "recurrent disease" subgroup analyses. The results presented here represent those provided by the manufacturer after clarification response,[20] as some amendments were made to the economic model. In the MS and the clarification response only ICERs are presented for the "treatment naïve" and "recurrent disease" subgroups, so the results presented in Table 12 have been obtained by the ERG re-running the model. As noted above, these results are based upon point estimates for parameters, and include several apparent errors identified by the ERG (see Section 5.3).

Table 12: Headline cost-effectiveness results presented by the manufacturer

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
All patients					
CYC	8.03	£95,819	-	-	-
RTX	8.19	£97,210	0.1628	£1,391	**£8,543.69**
Treatment naïve					
CYC	8.45	£81,327	-	-	-
RTX	8.53	£86,021	0.0851	£4,694	**£55,174.92**
Recurrent disease					
CYC	7.89	£100,699	-	-	-
RTX	7.98	£104,550	0.0896	£3,851	**£43,003.05**

In all population subgroups, the model suggests that the RTX treatment sequence is expected to produce more QALYs than the CYC sequence, albeit at a higher cost. The base case "all patients" analysis suggests that the incremental cost-effectiveness of the RTX sequence versus the CYC sequence is estimated to be around £8,544 per QALY gained. The base case "treatment naïve" analysis suggests that the incremental cost-effectiveness of the RTX sequence versus the CYC sequence is estimated to be around £55,175 per QALY gained, and the base case "recurrent disease" analysis suggests that the incremental cost-effectiveness of the RTX sequence versus the CYC sequence is estimated to be around £43,003 per QALY gained.

Probabilistic sensitivity analysis

The probabilistic version of the manufacturer's model produces considerably different results to the deterministic results. In the MS the probabilistic ICER is not stated, and in the economic model an "average" ICER calculated over 1000 simulations is generated. However this "average" is not appropriate, due to the influence of negative ICERs in some simulations. Instead, the ERG estimated the probabilistic ICER by dividing the average incremental costs over the 1,000 simulations (£687.50) by the average incremental QALYs (0.1715), generating an ICER of £4,008.22 for the "all patients" analysis. Through analysing the model in detail, the ERG has ascertained that the key reasons for the difference in the probabilistic result are the log normal distributions placed around the cost parameters in the model, the beta distributions placed around the relapse rates (which are arbitrarily estimated based upon standard deviations equal to 30% of the mean, which generates a wider distribution for the RTX relapse rate), and due to instability caused by the limited number of probabilistic samples. This is discussed further in Section 5.3.

It should be noted that the ERG could not replicate the exact PSA results reported in the MS or in the clarification response, but similar results were obtained. Whereas the manufacturer quotes probabilities of the RTX treatment sequence being cost-effectiveness of 61.7% and 64.6% for cost-effectiveness thresholds of £20,000 and £30,000 respectively (see MS p.172),[8] the figures obtained when the ERG re-ran the PSA using the model provided after clarification response were 62.2% and 71.2%. However, given the difference between the probabilistic and deterministic ICERs the ERG believes that these probabilistic results should be disregarded.

Simple sensitivity analysis

The simple sensitivity analysis presented within the MS (see MS p.171) and updated in the clarification response indicates that the base case "all patients" analysis was relatively insensitive to variations in the tested model input parameters with the exception of (i) the CYC remission rate, (ii) the uncontrolled disease utility (iii) the frequency of consultant visits in each health state and (iv) the reference cost applied to the consultant appointments in the "uncontrolled disease" health state. The

ICER for the RTX treatment sequence was markedly higher when the CYC remission rate was assumed to be the same as that for RTX (ICER=£24,656 per QALY gained); when the uncontrolled disease utility was assumed to be the same as the utility in the non-response health state (ICER=£14,404 per QALY gained); and when the frequency of consultant visits was reduced by 50% in each health state (ICER=£22,176 per QALY gained). Conversely, when the frequency of consultant visits in each health state was increased by 50% the RTX treatment sequence was estimated to dominate the CYC treatment sequence (lower costs and higher QALYs). The manufacturer's clarification response showed that when the unit cost applied to the outpatient appointments received while in the "uncontrolled disease" health state was reduced from £254.06 to £125.68, the ICER increased to £32,063 per QALY gained. This is a large change which demonstrates the impact of the resource use assumption for this health state. The model was sensitive to a slightly lesser extent to the assumed mean patient weight, and the proportion of patients that were assumed to receive IV rather than oral CYC treatment. The ICER for RTX versus CYC increased to £12,618 per QALY gained when the mean patient weight was increased to 87kg and reduced to £5,522 per QALY gained when 100% of patients were assumed to receive IV CYC in the control group.

Structural sensitivity analysis

The structural sensitivity analysis reported by the manufacturer in the MS and in the clarification response demonstrated that the results of the economic model were very sensitive to the sequences modelled. The results of these analyses are shown in Table 13.

Table 13: Structural sensitivity analysis cost-effectiveness results presented by the manufacturer

Option	Scenarios	ICER
Number of CYC courses	Base case: CYC group – 2 courses RTX group – 1 course	**£8,543.69**
	CYC group – 1 course RTX group – 0 courses	**£23,633.55**
	CYC group – 2 courses RTX group – 0 courses	**RTX dominated**
	CYC group – 2 courses RTX group – 2 courses	**RTX dominates**
	CYC group – 1 course RTX group – 1 course	**RTX dominates**
Which patients receive a second course of RTX?	Base case: Patients who initially do not achieve remission on RTX	**£8,543.69**
	No patients	**£7,197.34**
	Patients who achieve remission and then relapse	**£9,502.83**
	All patients	**£9,663.61**

In some scenarios, the RTX treatment sequence was dominated by the CYC treatment sequence, whereas in others RTX dominates CYC. The results demonstrate that if no patients initially treated with RTX are subsequently given CYC, the ICER for the RTX treatment sequence becomes much higher (less favourable). This appears logical because in these scenarios the CYC treatment sequence involves the same number or more lines of therapy compared to the RTX sequence – RTX is modelled as *replacing* CYC in the treatment pathway, rather than being *additional* to it. Hence in these scenarios patients have fewer opportunities to achieve remission in the RTX sequence than in the base case analysis. Conversely, in scenarios whereby the same number of courses of CYC are given in the RTX sequence as in the CYC sequence (indicating that two potential courses of RTX represent two additional lines of therapy, rather than one additional line of therapy as modelled in the base case), the ICER falls sharply (more favourable).

Assumptions around which patients will receive a second course of RTX have smaller impacts upon the ICER. Generally, increasing the proportion of patients who will receive a second course of RTX increases the ICER, and reducing the proportion of patients who will receive a second course of RTX decreases the ICER.

Based upon these analyses the ERG believes that there are two key questions relating to the potential cost-effectiveness of RTX:

1) Does the inclusion of RTX in the treatment sequence increase health benefits compared to the current treatment sequence?
2) If so, where is the most cost-effective place in the pathway to position RTX?

These questions are considered further in the ERG's analysis, presented in Section 6.1.3.

Subgroup analysis

The base case "treatment naïve" and "recurrent disease" subgroup analyses results are presented in Table 12. Table 14, demonstrates the sensitivity of the cost-effectiveness results to the assumed treatment sequences for the modelled subgroups.

Table 14: Structural sensitivity analysis cost-effectiveness results presented by the manufacturer

Subgroup	Scenarios	ICER
Treatment naïve	Base case: CYC group – 2 courses RTX group – 1 course	**£55,174.92**
	CYC group – 2 courses RTX group – 2 courses	**£1,273.92**
Recurrent disease	Base case: CYC group – 2 courses RTX group – 1 course	**£43,003.05**
	CYC group – 0 courses RTX group – 0 courses	**£12,556.29**

Given the sensitivities of the model results to the modelled sequences it is clearly of high importance to identify the most appropriate treatment sequences for each subgroup. This is discussed further in Section 5.3.

The manufacturer also provides a subgroup analysis for patients who are intolerant to CYC. In the comparator arm these patients receive best supportive care and transit immediately to the "uncontrolled disease" health state. In this analysis RTX dominates best supportive care.

5.3 Critique of the manufacturer's economic analysis

This section presents a critique of the manufacturer's economic model and associated analysis. This critique is set out according to four main sections:

(1) Adherence to NICE's Reference Case;
(2) Issues pertaining to the conceptual basis of the model and its structural assumptions;
(3) Issues relating to the evidence used to inform the model parameters;
(4) Technical issues relating to the implementation of the model.

5.3.1 Issues pertaining to the scope of the economic analysis and adherence to the NICE Reference Case

The manufacturer's model has been implemented generally in line with NICE's Reference Case (see Table 15). The economic analysis generally meets the scope issued by NICE.[19] Four deviations from the NICE scope warrant more detailed discussion: these relate to (i) the population, (ii) the specific indication for which RTX is modelled, (iii) the comparators included in the model and (iv) the outcomes included in the model. These are discussed below.

Table 15: Adherence to the NICE Reference Case

Element of health technology assessment	Reference Case	ERG comments
Defining the decision problem	The scope developed by the Institute	The scope of the economic analysis is generally in line with that developed by NICE.
Comparator	Therapies routinely used in the NHS, including technologies regarded as current best practice	Oral (2mg/kg/day) and IV (15mg/kg at 2-3 week intervals) CYC combined with corticosteroids followed by AZA (2mg/kg/day) (defined as standard care) are included in the economic analysis. Other treatments included in the scope (MTX and MMF) are not included in the model – the manufacturer argues they are not relevant.[20]
Perspective on costs	NHS and PSS	An NHS perspective was adopted which reflects costs over a lifetime time horizon.
Perspective on outcomes	All health effects on individuals	Health benefits for patients are measured and valued over a lifetime.
Type of economic evaluation	Cost-utility analysis	The economic analysis takes the form of a cost-utility analysis
Synthesis of evidence on outcomes	Based on systematic review	Systematic reviews were undertaken but little useful evidence identified (indirect comparisons were drawn). Response rates, relapse rates, adverse events and utility scores were largely based upon the RAVE trial, with additional assumptions required. Mortality rates and resource use were not based upon systematic review.
Measure of health effects	QALYs	Health outcomes are valued using QALYs.
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	SF-36 data were collected in the RAVE trial, and these were mapped to EQ-5D using a published regression equation.[66]
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	EQ-5D.
Discount rate	An annual rate of 3.5% on both costs and health effects	Costs and health outcomes are discounted at 3.5%.
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting is applied to QALY gains.

MS=Manufacturer's Submission; HRQoL=health-related quality of life; PSS=Personal Social Services; QALY=quality-adjusted life year.

(i) Population

The patient population addressed by the manufacturer's statement of the decision problem differs from that described in the final scope. Whereas the final scope intended for the appraisal to address all patients with ANCA-associated vasculitis, the MS only considers patients with "severe" MPA or GPA vasculitis. The rationale for this is that the anticipated RTX license will only cover this more limited population. In addition, the two pivotal RCTs upon which this appraisal rests, RAVE and RITUXVAS, only included patients with GPA and MPA AAV. Given the anticipated license and the main trial populations, it seems appropriate that the MS focuses upon "severe" GPA and MPA vasculitis. However, care must be taken with the definition of "severe" GPA and MPA vasculitis, since the definition of severity is not clear-cut, and the populations included in the RAVE and RITUXVAS trials differed. In RAVE, only patients with severe disease were recruited – these patients were defined as those with one or more of the major BVAS/WG items, or disease severe enough to require treatment with CYC. However, it is notable that the RAVE trial excluded patients with severe disease who required mechanical ventilation because of alveolar haemorrhage, and patients with a serum creatinine level of greater than 4.0mg/dL attributed to underlying AAV disease (see MS[8] p.35) – thus the clinical evidence submitted is not relevant for *all* patients with severe disease. The ERG believes that the exclusion criteria applied in RAVE mean that patients with severe renal disease were excluded, as well as patients with other life-threatening forms of the disease.

The EUVAS and BSR guidelines categorise severe ANCA-associated vasculitis as disease including renal or other vital organ failure, with serum creatinine of greater than 500μ/mol/litre (5.6mg/dL). Hence it seems that these patients would actually have been excluded from the RAVE trial.[9,17] Under these guidelines, the RAVE definition of severe disease appears closer to what is classified as "generalised" disease – where vital organ function is threatened and serum creatinine is less than 500μ/mol/litre (5.6mg/dL).

There is an additional question as to whether the population included in the RAVE trial, which forms the basis of the manufacturer's economic evaluation, included a population relevant to the UK population. The manufacturer states that the RAVE study recruited mainly in the United States and the Netherlands, and that the average body surface area (BSA) of included patients was 1.998m^2, with an average weight of 87kg (although other evidence submitted by the manufacturer suggested that the average weight was actually 85.1kg).[21] The manufacturer states that this weight and BSA is not reflective of UK patients, and instead an average BSA of 1.79m^2 and an average weight of 67kg was assumed based upon estimates for UK cancer patients (see MS[8] p.138). As will be discussed in Section 5.3.3 it appears that this is likely to underestimate the typical weight of UK patients with ANCA-associated vasculitis. This is important due to the weight and BSA-related doses of several of the treatments included in the manufacturer's economic model. Overall, the ERG believes that the

clinical trial results and those of the economic model may not be generaliseable to *all* patients with severe GPA or MPA vasculitis. In addition, while the results from the RAVE trial may be relevant for a heavier population than is typically seen in the UK, the base case model results may underestimate general patient weight and BSA. The ERG has undertaken further analysis investigating this issue, reported in Section 6.

(ii) Indication

The specific indication described in the MS differs slightly from that described in the final scope developed by NICE due to its anticipated license. The final scope did not specify whether the intervention would be considered for induction, maintenance, or relapse therapy. However the manufacturer anticipates that the intervention will be licensed only for induction therapy and so only this indication is considered in the MS. This is despite the fact that the RAVE trial included the treatment of flares with RTX, and other clinical trials have studied maintenance treatment with RTX. This causes some problems for the decision problem faced in the appraisal, because the ERG's clinical advisors state that it is very unlikely that in reality patients who achieve remission through treatment with RTX would go on to receive no maintenance therapy, either with RTX or another treatment (such as AZA) – yet it is this treatment pathway that the manufacturer models in their economic model. The clinical advisors to the ERG suggest that patients treated successfully with RTX will regain their B-cells over time and are therefore likely to relapse, suggesting that without maintenance therapy the disease may be badly controlled. Given that the manufacturer's economic model is based primarily on the RAVE trial which included no maintenance treatment in the RTX arm, and that the anticipated license will not include RTX maintenance therapy, this represents an important issue. Given the anticipated license, and advice provided by clinical advisors to the ERG, it would appear appropriate to assume that patients who achieve remission on RTX then receive maintenance therapy with AZA – representing the maintenance treatment received in the CYC arm of the RAVE trial – or MTX. The manufacturer did not consider such a treatment pathway in their economic analysis.

It is also important to note that in the manufacturer's economic model it is assumed that patients who initially do not respond to RTX treatment are immediately given a second course of RTX, offering a further chance of achieving remission. Such retreatment has not been studied in any of the evidence submitted by the manufacturer, but it appears not to be ruled out by the anticipated license. The manufacturer states that they modelled this sequence based upon expert opinion. However advice received from the ERG's clinical advisors is contrary to this – suggesting that there is no evidence in favour of this and that it would be more likely that an alternative therapy would be tried. Hence it is not clear that the treatment pathways modelled in the manufacturer's base case are appropriate – however the model has the capacity to alter the proportion of patients that go on to receive RTX

retreatment (non-responders [as in the base case], relapsers, all patients, or no patients). The ERG has investigated alternative treatment sequences in Section 6.

(iii) Comparators

The manufacturer did not include MTX or MMF as comparators despite these being included as comparators in the NICE scope.[19] In response to a question about comparators in the clarification process, the manufacturer stated (see Clarification Response, page 14):[20]

"The clinical advice we received outlined clearly that for patients to be considered for RTX, the only relevant comparator technology is CYC. The clinician we spoke with said that whilst patients with milder forms of AAV may receive mycophenolate or methotrexate, these patients are generally not considered to be candidates for RTX and indeed do not fall within the scope of this appraisal."

While the ERG accepts that CYC represents the most relevant comparator for the population who would be considered for RTX, we contend that MTX and MMF have relevant parts to play in treatment sequences that should have been considered in the economic model. In addition, the ERG notes that early results of the MYCYC trial recently published in abstract form suggest that MMF has demonstrated very similar efficacy to CYC in a relevant, moderately severe, patient group.[22]

Of particular importance is that the manufacturer's model assumes that after receiving two courses of CYC the current standard of care is for patients to receive only "best supportive care" and reside in an "uncontrolled disease" health state – a health state in which the health related quality of life is poor (the utility score is 0.67, compared to 0.75 in the "non-response" state and 0.83 in the "remission" health state) and resource use is high (one "best supportive care" outpatient appointment every 1.5 weeks). In the manufacturer's model, patients in the "standard care" CYC group spend 70.7% of their discounted mean life expectancy in this health state (8.0% is spent in the "remission" state and 21.4% is spent in the "non-response" state), compared to 63.2% in the RTX group (9.3% is spent in the "remission" state and 27.6% is spent in the "non-response" state). However, the ERG's clinical advisors suggest that it is very rare for patients with severe GPA or MPA vasculitis to reside in such a state – usually a treatment strategy can be identified that offers some control of the disease, and this may involve treatments such as MTX, MMF, higher dose AZA or co-trimoxazole, intravenous immunoglobulin, infliximab, leflumonide, abatacept, calcineurin inhibitors or combinations of these. Such a health state may be described as low grade "grumbling" disease.[71] The ERG believes that the failure to allow for multiple treatments that are likely to offer some degree of disease control within the manufacturer's economic model may distort the cost-effectiveness results and exaggerate the economic attractiveness of RTX. However, considering that subsequent induction treatment with alternative interventions is excluded from both arms of the manufacturer's model the impact may be

relatively minor. The ERG suggests that what is likely to be most important is a consideration of whether RTX represents an additional treatment in the treatment pathway, or whether it replaces another treatment. The manufacturer's model implicitly assumes that RTX offers an additional line of therapy, since two courses of RTX may be received followed by one course of CYC – 3 lines of treatment in total, compared to the 2 courses of CYC modelled in the control arm. While the ERG believes that it may not be appropriate to assume that patients will receive re-treatment with RTX following RTX failure, clinical advice suggests that use of RTX may not necessarily reduce lifetime use of CYC (apart from in circumstances where a RTX-induced remission does not relapse) – patients may still receive CYC if RTX does not work for them. Hence considering RTX as an additional treatment in the pathway seems reasonable. We attempt to address treatment sequencing issues in Section 6.

In addition, even for the subgroup analysis of patients who are intolerant to CYC the administration of other treatments was not modelled – it is instead assumed that patients on "standard care" immediately enter the "uncontrolled disease" health state. The ERG believes that this leads the cost-effectiveness case for RTX to be over-stated in this subgroup. However, we accept that alternative treatments have not been demonstrated to be as effective as CYC in this patient group (although the NORAM study has demonstrated similar remission rates comparing CYC with MTX in less "severe" patients (but generally poorer relapse-free survival),[27] and early results from the MYCYC suggest similar may be true for MMF in moderate and "severe" patients.[72,22]

(iv) Outcomes

The manufacturer has not included all of the outcomes listed in the scope in their economic evaluation. Although the main cost-effective outcome is in the form of an incremental cost per QALY gain, as required, the ERG believe that the model may be too simplistic to satisfactorily model the disease pathway, and that the model structure means that some important clinical outcomes are excluded. Different severities of relapse and different grades of remission were not modelled by the manufacturer. Clinical advice received by the ERG suggests that treatment options and the subsequent disease pathway depend critically upon whether a minor or major relapse is experienced. Data on relapse severity were available from the RAVE trial and are presented in the clinical section of the MS (see p.68), but different relapse severities were not included in the economic model. In their response to clarification questions the manufacturer states (see manufacturer's Clarification Response, p.11):

"Our economic model defines all relapses as necessitating treatment with CYC or RTX (depending on which arm patients are in). We chose to use this assumption because, although in reality patient progression in AAV is a complicated path with a variety of events possible, our main trial does not

capture this complexity. We felt it was more appropriate to develop a simple model whose parameters can be varied in sensitivity analysis.

The clinical expert we consulted said that with close monitoring, most relapses are identified while they are still mild, and if left untreated most of these will subsequently develop into major relapses. This means that in clinical practice, although some patients experiencing a minor relapse may just have their steroid dose increased, almost all will ultimately have a more serious relapse justifying treatment (or retreatment) with CYC/RTX.

Our clinical expert considers retreatment with CYC or RTX at the first sign of relapse. For example, if [a] patient relapses whilst on CYC, treatment with RTX would be considered immediately to avoid irreversible harm potentially caused by waiting for the relapse to develop into a major one."

Hence, in their economic model, the manufacturer assumes that all relapses lead to immediate retreatment with CYC or RTX. The manufacturer rationalises this by suggesting that almost all minor relapses will lead to major relapse requiring retreatment. However, it is the understanding of the ERG (based upon clinical advice) that minor relapses may be well controlled through relatively minor treatment using, for example, an increase in corticosteroid dose – re-treatment with CYC of RTX may not be required. In particular, the ERG's clinical advisors wholly disagree with the statement that almost all patients that experience a minor relapse will go on to experience a major relapse – if the minor relapse is treated appropriately (not with re-induction treatment), only a minority would go on to experience a major relapse.

The ERG's clinical advisors also suggest that the failure to model different levels of treatment response may be important. The manufacturer's model defines four health states – "complete remission", "non-remission", "uncontrolled disease" and "death". Complete remission is defined by a BVAS/WG score of 0 and a successful taper of glucocorticoid therapy at 6-months after randomisation, as defined in the RAVE trial. However, our clinical experts suggest that a proportion of patients are likely to achieve a BVAS/WG score of 0 but may still be receiving glucocorticoid treatment at 6-months. Data on such patients are presented in the clinical section of the MS (for instance, see Table 15 on p.68 of the MS).[8] These patients are likely to have different levels of quality of life and resource use compared to patients who do not achieve a BVAS/WG score of 0 (they are unlikely to require re-treatment with CYC or RTX until a subsequent major relapse), but these patients are classified as non-responders in the economic model and go on to receive a second course of treatment. The difference in the proportion of patients who achieved complete remission and remission without a completed glucocorticoid taper was similar in the RTX and CYC groups in the RAVE trial, and therefore the relative impact of not including a health state for these "non-complete

responders" in the economic model may be minor – however the knock-on impacts on subsequent treatments, costs and effectiveness in the economic model is unknown.

Overall, the impact on the cost-effectiveness results of not capturing the true disease and treatment pathway in the economic model are unknown. However, the cumulative effect of the assumptions made by the manufacturer is likely to lead to patients entering the "uncontrolled disease" health state inappropriately quickly in both treatment groups. Given the low HRQoL and high cost associated with this state it is likely that the bias will be primarily against the treatment group which enters this state most quickly – hence the ERG suggest that the simple nature of the economic model may result in bias against CYC, though the magnitude of this bias may be relatively minor.

5.3.2 Issues pertaining to the conceptual basis of the model and its structure

The manufacturer states that they asked one clinical expert to comment on the model structure. They also state that the "uncontrolled disease" health state was informed by expert opinion and that the treatment sequences modelled were based upon clinical advice – in particular that CYC treatment is limited to two courses, and that RTX re-treatment occurs immediately in patients who initially fail to respond to a first course of RTX treatment. The ERG believes that there are important issues surrounding the treatment sequences modelled and the health states included.

As stated in Section 5.3.1, the ERG believes that a more appropriate economic model would have included a health state for a non-complete remission (i.e. where corticosteroids and other less immunosuppressive treatments such as high dose co-trimoxazole are still used to control the disease) and would model relapses of varied severity. However it is beyond the remit of the ERG to re-build the economic model including these factors. Perhaps more importantly, the ERG believes that the "uncontrolled disease" health state as modelled by the manufacturer is unrealistic, and does not represent a state of health commonly seen in ANCA-associated vasculitis patients (certainly not one that patients spend the majority of their lives in). The ERG's clinical advisors suggest that a more common health state is one in which the most effective induction treatments have been used, but some other treatment or combination of treatments is utilised in order to afford patients a reasonable level of disease control. The ERG does not believe that the level of HRQoL associated with this low-grade "grumbling disease" health state is significantly and substantially worse than the HRQoL in the "non-remission" health state, particularly if steroid doses can be minimised or withdrawn. Although treatment would undoubtedly be received in this state, the assumption that specialist palliative care is received at hospital outpatient appointments once every 1.5 weeks for the remainder of their life appears to be a substantial over-estimate which is not supported by evidence and does not match clinical practice in the NHS. Clinical advice received by the ERG suggests that it would be more

appropriate to assume that patients in this health state continue to receive maintenance treatment, and that outpatient appointments occur each month initially, followed by less frequent visits over time.

The main analyses presented by the manufacturer are for "all patients" – this includes a mixture of treatment naïve patients and patients with recurrent disease. As well as this, the manufacturer presents subgroup analyses for treatment naïve patients and patients with recurrent disease. The ERG believes that these are important and relevant analyses, and may help inform optimal treatment strategies for these different patient subgroups. The key disadvantage of these analyses is that each is based on only approximately half of the trial data from RAVE (because approximately half of the patients in RAVE were treatment naïve, and half had recurrent disease). Hence, these analyses are subject to a greater degree of uncertainty. Because of the importance of these analyses in the remainder of this section, the structure of the model is critiqued separately for the "all patients" analysis, the "treatment naïve" analysis, and the "recurrent disease" analysis. A section on patients intolerant to CYC is also included. In addition to the discussion on relevant sequences of induction treatments, it is important to note that, based upon clinical advice received by the ERG, it is very unlikely that patients who achieve remission will receive no maintenance therapy. Given this, and that RTX is not licensed for maintenance therapy, it seems most appropriate to assume that patients who achieve remission after taking RTX receive AZA maintenance therapy. This is not what was observed in the RAVE trial, and has implications for the relapse rates used in the model. However, the ERG believes that in the interests of running scenarios that reflect realistic treatment pathways it is useful to run scenarios in which RTX patients receive AZA maintenance therapy – the assumptions around this are discussed in Sections 5.3.3 and 6.1.2. Note that no analyses are run where RTX is given as maintenance therapy – although there is evidence of the effectiveness of RTX in this setting, it is not expected to be included in the license.

All patients

In the manufacturer's base case "all patients" analysis the standard care arm of the economic model involves a maximum of two courses of CYC. Upon relapse or non-response to the second course of CYC patients enter the "uncontrolled disease" health state. In the experimental arm of the economic model patients receive a maximum of two courses of RTX followed by one course of CYC. The second RTX course is only received by patients who do not respond to initial RTX treatment. Upon relapse or non-response to the course of CYC patients enter the "uncontrolled disease" health state.

In their clarification response, the manufacturer states that their assumption that two courses of CYC represents standard care was based upon clinical advice on the number of courses of CYC IV pulse therapy that would typically be given.[20] Clinical advice received by the ERG suggests that this assumption is questionable. The ERG understands that there is no consensus on a specific lifetime

maximum dose of CYC that should not be exceeded – however, a cumulative dose of 20g-30g appears to provide a range that should not be exceeded. One complete course of oral CYC (2mg/kg/day for 6 months) is associated with a total dose of 24,471.75mg (24.47g) for a patient weighing 67kg (as in the base case analysis). The total dose would be 31,046.25mg (31.05g) for a patient weighing 85kg (the mean weight in the RAVE trial). One complete course of IV CYC (15mg/kg ten times over a 6 month period) is associated with a total dose of 10,050mg (10.05g) for an 67kg patient, and 12,750mg (12.75g) for an 85kg patient. Hence the manufacturer's assumption that two courses of CYC would be received by patients in the standard care arm would only appear reasonable if those courses were IV CYC. This is in line with the expert opinion received by the manufacturer, but does not reflect the assumptions employed within the manufacturer's economic model, as the model assumes that 28% of CYC treatment is given as a 6-month oral course. If 6-months oral CYC is the comparator, it would seem appropriate to only allow one course of treatment. However, it is important to note that while the manufacturer's model assumes that oral CYC is given for 6 months, in the RAVE trial treatment was for 3-6 months and in fact the cumulative dose was 15,234.02mg (standard deviation 7,199.724), which, given the mean weight in the CYC group in RAVE of 87.88kg, is substantially below the full oral dose (32,098mg) had 6 months of treatment been received. In these circumstances it seems possible that two doses of oral CYC may be given.

Clinical advice received by the ERG suggests that it is unlikely that a patient who did not respond to an initial course of RTX treatment would receive a second course, due to the lack of evidence for this. It would be more likely that such a patient would move onto an alternative therapy. It *may* be clinically reasonable to expect that a patient who previously responded well to RTX would receive it again upon disease relapse, but there is currently no published evidence for such use.

RTX appears to represent an additional treatment option, rather than replacing an existing treatment. Hence, in the event of relapse there appears to be no clinical reason why patients who initially start treatment on RTX cannot end up receiving the same number of courses of CYC as those patients who begin on CYC treatment. It is also important to consider that there is no reason why RTX must only be considered as the *first* induction treatment received in the economic model. It is relevant to consider the relative cost-effectiveness of RTX used before and after CYC in the treatment pathway. Indeed, the NHS Commissioning Board only recommend the use of RTX as an initial induction agent in newly diagnosed patients where avoiding the use of CYC is desirable.[18] Otherwise, RTX is recommended as a remission induction agent when CYC has not been effective, or at the time of first relapse.[18] The cost-effectiveness of RTX for different patient groups (treatment naïve and recurrent disease) can be assessed using subgroup analyses as undertaken by the manufacturer. However, the cost-effectiveness of different orders of treatment could also be assessed within the "all patients" analysis.

It is difficult to determine the most appropriate number of courses of treatment to assume for the "all patients" analysis – particularly because half of the patients included in the analysis have recurrent disease and are therefore likely to have already received CYC (limiting the scope for them to receive subsequent courses). This demonstrates why the "treatment naïve" and "recurrent disease" subgroup analyses are so important – because appropriate treatment sequences can be defined much more easily. However, the ERG suggest that for this "all patients" analysis the following treatment sequences are relevant and should be assessed (with "supportive care" representing the low grade "grumbling" disease that the ERG believe is more realistic than the "uncontrolled disease" state; and where oral CYC is assumed to be received for 3-6 months as in the RAVE trial, rather than for the full 6 months):

i) CYC (IV/oral) → CYC (IV/oral) → Supportive care
ii) CYC (IV/oral) → RTX → CYC (IV/oral) → Supportive care
iii) CYC (IV/oral) → CYC (IV/oral) → RTX → Supportive care
iv) RTX → CYC (IV/oral) → CYC (IV/oral) → Supportive care

Treatment naïve subgroup

In the manufacturer's base case "treatment naïve" analysis, the treatment sequences received in the standard care and experimental arms of the economic model are identical to those received in the "all patients" analysis. In the standard care arm, a maximum of two courses of CYC are given. Upon relapse or non-response to the second course of CYC patients enter the "uncontrolled disease" health state. In the experimental arm of the economic model patients receive a maximum of two courses of RTX followed by one course of CYC. The second RTX course is only received by patients who do not respond to initial RTX treatment. Upon relapse or non-response to the course of CYC patients enter the "uncontrolled disease" health state.

Based upon cumulative CYC doses, the ERG believe that treatment naïve patients in the standard care arm of the economic model would appropriately receive two courses of IV CYC, or two courses of oral CYC if treatment was for 3-6 months, rather than a full 6-month course. A similar total number of CYC courses could be received either before or after one course of RTX. The ERG suggests that the same treatment sequences as specified for the "all patients" analysis would be relevant for consideration within the "treatment naïve" analysis (again, with "supportive care" representing the low grade "grumbling" disease that the ERG believe is more realistic than the "uncontrolled disease" state; and where oral CYC is assumed to be received for 3-6 months as in the RAVE trial, rather than for the full 6 months):

i) CYC (IV/oral) → CYC (IV/oral) → Supportive care

ii) CYC (IV/oral) → RTX → CYC (IV/oral) → Supportive care

iii) CYC (IV/oral) → CYC (IV/oral) → RTX → Supportive care

iv) RTX → CYC (IV) → CYC (IV) → Supportive care

For the "treatment naïve" analysis we can be more certain that these represent appropriate treatment sequences, since we know that all patients being considered are treatment naïve. However, data for analyses based upon this subgroup are more scarce, since only half of the patients recruited to RAVE were treatment naïve.

Recurrent disease subgroup

In the manufacturer's base case "recurrent disease" analysis, the treatment sequences received in the standard care and experimental arms of the economic model are identical to those received in the "all patients" analysis. In the standard care arm a maximum of two courses of CYC are given. Upon relapse or non-response to the second course of CYC patients enter the "uncontrolled disease" health state. In the experimental arm of the economic model patients receive a maximum of two courses of RTX followed by one course of CYC. The second RTX course is only received by patients who do not respond to initial RTX treatment. Upon relapse or non-response to the course of CYC patients enter the "uncontrolled disease" health state.

Based upon cumulative CYC doses, the ERG believe that patients with recurrent disease would appropriately receive one course of IV CYC, or one course of 3-6 months oral CYC (rather than a complete 6-month course). Potentially zero courses of CYC would be given if previous treatment had been with a complete 6-month course of oral CYC. Similar treatment options would be available in a RTX treatment sequence, given either before or after one course of RTX. In this subgroup it is assumed that previous induction treatment was with CYC; in practice, if disease is moderately severe or mild, other agents such as MMF may have been used. In these situations the relevant treatment sequences would be those specified for the "treatment naïve" subgroup – reflecting patients who are CYC-naïve. The ERG suggests that the following treatment sequences would be relevant for consideration within the "recurrent disease" analysis, for patients previously treated with CYC (again, with "supportive care" representing the low grade "grumbling" disease that the ERG believe is more realistic than the "uncontrolled disease" state; and where oral CYC is assumed to be received for 3-6 months as in the RAVE trial, rather than for the full 6-months):

i) CYC (IV/oral) → Supportive care

ii) Supportive care

iii) CYC (IV/oral) → RTX → Supportive care

iv) RTX → CYC (IV/oral) → Supportive care

v) RTX → Supportive care

For the "recurrent disease" analysis we can be more certain that these represent appropriate treatment sequences, since we know that all patients being considered have recurrent disease. However, data for analyses based upon this subgroup are more scarce, since only half of the patients recruited to RAVE had recurrent disease.

CYC intolerant subgroup

The manufacturer presents a subgroup analysis for patients who are intolerant to CYC. This is problematic, because patients that were intolerant of CYC were excluded from the RAVE clinical trial and basing an analysis of this subgroup on any data from RAVE therefore seems highly suspect. However, clinical advice received by the ERG suggests that there is no reason to expect that the clinical effectiveness of RTX would differ in patients who are intolerant of CYC compared to patients who are tolerant of CYC. Therefore, using response and relapse rates for RTX from RAVE seems reasonable. However, the treatment sequence assumed for patients in the standard of care arm of the model seems unrealistic and inappropriate – it is assumed that these patients immediately enter the "uncontrolled disease" health state, incurring high specialist palliative care costs every 1.5 weeks and experiencing a low HRQoL. This assumption does not seem to reflect usual care in the NHS. It would be more appropriate to draw a comparison to an alternative treatment, such as MTX, MMF or leflunomide for this subgroup analysis, although the ERG accepts that data for the use of these treatments in patients who have "severe" ANCA-associated vasculitis is currently scarce. The ERG believes that allocating a higher utility score to the "uncontrolled disease" health state, more in line with the more realistic low-grade "grumbling" disease, and assuming lower resource use, in line with advice from our clinical experts, would allow a more appropriate analysis of the cost-effectiveness of RTX in this subgroup.

5.3.3 Issues relating to the evidence used to inform the model parameters

Remission rates

Remission rates are estimated primarily based upon the RAVE trial, although for some parts of the modelled treatment sequences data are not available and assumptions are made. The ERG's main criticism of the remission rates utilised is that an arbitrary assumption is made that the relative decrement in effectiveness of a second course of RTX is the same as the relative decrement in effectiveness of a second course of CYC – based upon the CYC remission rate in treatment naïve patients and patients with recurrent disease. Currently no data exist on the effectiveness of a second course of RTX and the ERG questions the assumption that non-responders immediately receive a second course.

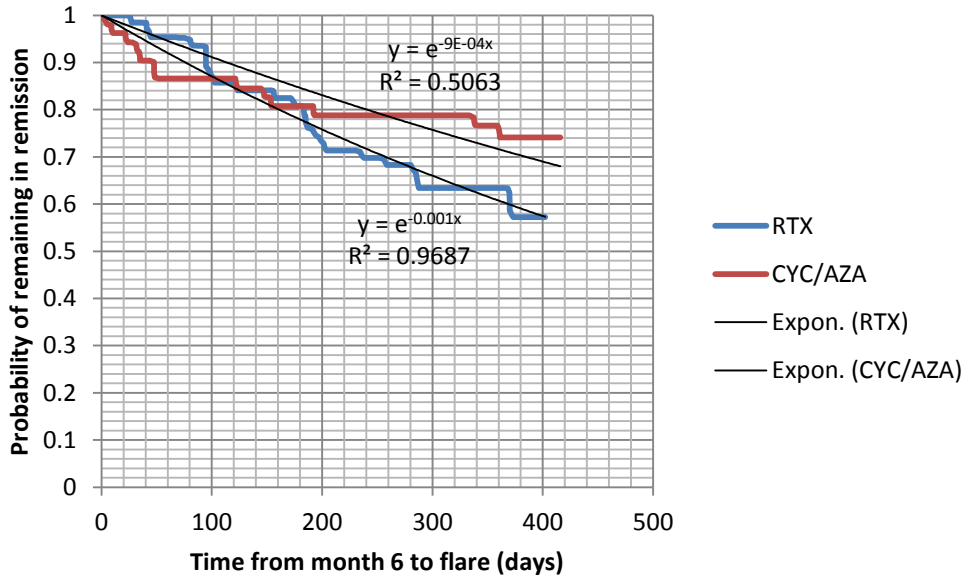
In the "all patients" analysis data on the effectiveness of a first course of CYC or RTX are taken from the remission rate in all patients randomised to the RAVE trial, by treatment group (subject to worst case imputation for missing data, which was minimal). The effectiveness of a second course (or the first course in the "recurrent disease" subgroup analysis) is based upon the remission rate in patients with recurrent disease who were included in the RAVE trial. While this seems reasonable, the ERG believes that this approach confuses the analysis somewhat, since some of the patients informing the first course remission rate in the economic model will have received prior treatment and this appears to impact upon the effectiveness of both CYC and RTX. Although based upon fewer data, the ERG believes that a more appropriate analysis would organise the model in a more structured way, based upon disease stage rather than course of treatment. Patients would enter the model treatment naïve, and would receive the remission rate from treatment naïve patients from RAVE. Following subsequent relapse or non-response patients would then have recurrent disease and would receive the remission rate from patients with recurrent disease from RAVE. This is similar to the "treatment naïve" subgroup analyses run by the manufacturer, although the ERG would suggest other alterations to these analyses with regard to other model parameters, as discussed below.

For an analysis of a "recurrent disease" subgroup the manufacturer has appropriately used remission rates for the first course of treatment based upon patients with recurrent disease included in the RAVE trial. However, for the second course of CYC treatment a reference to Hoffman *et al.* (1992)[15] was given, but it is not clear how data from this study were used to estimate the appropriate remission rate. However, given the concerns about administering multiple courses of CYC over a lifetime this may be of little importance, because it seems more likely that most patients with recurrent disease would only receive a maximum of one further course of CYC, with relatively few having had initial treatment with regimes not based on CYC.

Relapse rates

The relapse rates used in the manufacturer's model were derived from the RAVE trial. Exponential models were fitted to data from patients who had experienced complete remission at 6-months in order to estimate the time-to-event for relapse (see MS p.136).[8] In their clarification response the manufacturer states that individual patient level data were not used to fit these models, instead they were fitted using summary data. The ERG believes that these rates have been estimated poorly. Figure 2 presents Figure 14 from the MS.[8]

Figure 2: Probability of remaining in remission (Figure 14 taken from MS)



Although the precise method used by the manufacturer to fit the exponential models is not outlined in the MS or the manufacturer's clarification response, the fact that patient-level data were not used to fit the exponential models means that the relative number of events observed at different time points is unlikely to have been taken into account. This means that events that occur towards the end of follow-up, when numbers at risk are smaller (presumably, though data have not been provided on numbers at risk), may receive inappropriate weight in the model fitting process.[73,74] In addition, the use of summary data rather than patient-level data precludes the use of model fit statistics such as Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) – the use of measures of explained variation such as R^2 are not recommended for use with time-to-event data.[75]

Ideally, log-cumulative hazard plots or similar would be used to justify the use of the exponential models, but these are not presented by the manufacturer. Figure 2 shows that the Kaplan-Meier relapse curves for RTX and CYC (CYC/AZA) cross, indicating that the proportional hazards assumption does not hold, and that applying constant relapse rates over time to each treatment group is unlikely to be appropriate. Given that the Kaplan-Meier curve for CYC appears to flatten over time it is possible that the relapse rate for the CYC group is over-estimated. It appears highly likely that an alternative parametric model (such as a Weibull, Gompertz, log normal, log-logistic etc.) would have provided a better fit to the relapse data than an exponential. The ERG assumes that the exponential model was chosen due to the ease with which the resulting constant rates can be applied within a standard Markov model structure. The ERG suggests that this may mean that the standard Markov model structure is inappropriate in this case.

A similar approach was taken in order to estimate relapse rates for the "treatment naïve" and "recurrent disease" subgroups. However, diagrams and statistics have not been provided by the manufacturer in order to allow the ERG to assess the fit of the exponential models used. In their clarification response the manufacturer states (see manufacturer's Clarification Response,[20] p.12-13):

"The fit from the exponential curves was based on summary statistics, not patient level data. This, as well as the low patient numbers in each arm (n=49), led to less precise estimates for relapse rates in subgroups. In particular, we note a counterintuitive result in the "relapsing patients" group (i.e. patients who previously received and relapsed on CYC). Specifically, the RTX arm was associated with high rates of relapse compared to the CYC arm; we would have expected the reverse to be true."

Given the data provided it is not possible for the ERG to assess the relative fit of the exponential models fitted to the subgroup relapse data. However, the ERG believes that the higher relapse rate observed in the "relapsing patients" (recurrent disease) subgroup may be intuitive, given that these patients had a history of relapsing and in the RAVE trial no maintenance therapy was provided in the RTX group, whereas it was provided in the CYC group.

Clinical advice received by the ERG suggests that the relapse rates used in the manufacturer's economic model appear unrealistically high. The rates used in the model are such that 1 year after achieving complete remission approximately 30% of patients have relapsed. Approximately 50% have relapsed after 2 years and approximately 90% have relapsed after 7 years. In a study comparing CYC to AZA for maintaining remission in 155 ANCA-associated vasculitis patients with similar severity to those included in the RAVE trial approximately 15% of patients who achieved remission experienced relapse at 18 months, with approximately half of these representing severe relapses.[12] A recently published abstract presenting the results of a prospective study testing the use of RTX compared to AZA for maintaining remission in 117 patients with ANCA-associated vasculitis reported major relapse rates of 25% and 5% in the AZA and RTX groups respectively, at 28 months.[76]

In addition to this, clinical advice received by the ERG suggests that only severe relapses would result in re-induction treatment. Limited relapses can usually be treated more moderately, for example with an increase in steroid dose. Although such relapses would prevent a patient from being defined as being in a "complete remission" health state, they would be unlikely to lead to re-induction treatment. As previously stated, ideally the manufacturer's model should include a health state for patients in remission who require steroids to maintain full control of their disease. However, in the absence of this, the ERG believes that it would be most reasonable to assume that only severe relapses lead to re-induction treatment – and thus patients who experience a limited relapse remain in the "complete remission" health state. In their response to the ERGs clarification questions, the manufacture

Superseded – see erratum

r contended that it is appropriate to assume that *all* relapses lead to re-induction treatment (Clarification Response,[20] p.11):

"Our economic model defines all relapses as necessitating treatment with CYC or RTX (depending on which arm patients are in). We chose to use this assumption because, although in reality patient progression in AAV is a complicated path with a variety of events possible, our main trial does not capture this complexity. We felt it was more appropriate to develop a simple model whose parameters can be varied in sensitivity analysis.

The clinical expert we consulted said that with close monitoring, most relapses are identified while they are still mild, and if left untreated most of these will subsequently develop into major relapses. This means that in clinical practice, although some patients experiencing a minor relapse may just have their steroid dose increased, almost all will ultimately have a more serious relapse justifying treatment (or retreatment) with CYC/RTX.

Our clinical expert considers retreatment with CYC or RTX at the first sign of relapse. For example, if patient relapses whilst on CYC, treatment with RTX would be considered immediately to avoid irreversible harm potentially caused by waiting for the relapse to develop into a major one.

Our economic model has not modelled different categories of relapse. Instead we assume that the majority of relapses will lead to retreatment with CYC or RTX (depending on arm). We acknowledge that some patients may not require retreatment with CYC or RTX immediately, but as we have no way of determining which minor relapses might eventually progress to major ones, nor what time scale such progression might happen on, we have not factored this distinction into our analysis."

This argument is contrary to the clinical advice received from the ERG's clinical experts. However it is noted that *"most relapses are identified while they are still mild, and if untreated most of these will subsequently develop into major relapses"*. This appears to be key in the manufacturer's argument in favour of assuming that all relapses lead to re-induction therapy. However, the ERG contends that mild relapses are likely to be treated, and therefore subsequent major relapse and re-induction treatment may be avoided. Hence, the ERG believes that only major relapses should be assumed to lead to re-induction treatment in the manufacturer's model.

Data from the RAVE trial provided by the manufacturer show that 16 severe flares were observed in the RTX group in 838.1 participant months, compared to 12 severe flares in 801.4 participant months in the CYC group. This is associated with monthly relapse rates of 0.019 and 0.015 in the RTX and CYC groups respectively, which are equivalent to 6-month probabilities of relapse of 10.8% and 8.6%

Superseded – see erratum

respectively. Applying these rates constantly over time leads to approximately 17-21% relapsing after 1 year and 30-36% relapsing after 2 years. These rates appear to be closer (though still higher) to relapse rates reported in other relevant studies,[12,76] and more closely reflect the expectations of the ERGs clinical experts. Such data are not available to the ERG for the "treatment naïve" and "recurrent disease" subgroups. The ERG anticipates that such evidence would involve very low event numbers and therefore would be highly uncertain. Therefore it may be preferable to assume similar relapse rates – given appropriate maintenance therapy – in these two groups. Clinical advice received by the ERG suggests that this would not be an inappropriate assumption, although it is possible that relapse rates may be higher for patients who have previous relapsed. Clinical advice received by the ERG also suggests that under an assumption that patients induced into remission by RTX received AZA maintenance therapy, assuming equal relapse rates in these patients and those induced by CYC (who subsequently receive AZA) would be reasonable.

Mortality rates

In the MS, it is assumed that the "base" SMR applied in the economic model is applicable to patients in the "non-remission" health state and an arbitrary ±10% is applied to this for patients in the "remission" and "uncontrolled disease" health states respectively. In their response to clarification questions the manufacturer noted that expert clinical opinion supported the notion that patients in clinical remission would have a better prognosis than patients with poor disease control (Clarification Response,[20] p.34). However, it is worthy of note that this view does not appear to be universal: in a 2012 paper clinical expert David Jayne states *"Relapse does not increase the risk of end-stage renal disease or death"*.[77] The Lane *et al.* study is small in size and dated – however, given that life expectancy predicted by the manufacturer's model seems reasonable and that the treatment of choice for vasculitis (CYC) has not changed substantially in recent years, the applied mortality rates based upon the Lane *et al.* (2005) study seem reasonable. It is debatable whether different rates should be applied to the different health states (particularly for the "uncontrolled disease" and "non-response" health states), but these differences lead to very marginal benefits for RTX and do not represent key drivers of the cost-effectiveness results.

Adverse events

AEs have relatively little impact in the manufacturer's model. Although the ERG has identified problems with the way that these have been incorporated in the economic model (inconsistent utility score data, potentially inappropriate assumed durations of events) these are very unlikely to be key drivers within the economic model. An issue which is potentially more important is that malignancies were not included in the economic model. However, in their response to clarification questions asked by the ERG the manufacturer stated that only two instances of malignancies were reported in the

Superseded – see erratum

RAVE trial, and neither could be reliably related to RTX (see manufacturer's Clarification Response,[20] p.37).

Utility scores

The methods used by the manufacturer to estimate utility scores for the "complete remission" and "non-response" health states appear to be appropriate. It is not clear exactly which data from RAVE were used to inform these estimates – the ERG assumes that utilities for "remission" were estimated based upon data from patients who achieved complete remission at 6-months, but it is unclear whether utilities for "non-remission" included only data from non-responders at 6-months, or also included data from all patients at baseline. Given the clinical advice received by the ERG – that the "uncontrolled disease" health state is very rare and that a much more common non-remission health state could be described as low-grade "grumbling" disease – it appears that the utility associated with the "uncontrolled disease" health state may be underestimated. If this health state were replaced with a "grumbling disease" health state a higher utility score is likely to be appropriate. This may be marginally lower than the utility score experienced in the "non-remission" health state, in which patients are receiving CYC or RTX as induction treatment, but seems unlikely to be substantially lower.

Drug costs

Generally the costs of drugs included in the manufacturer's economic model appear to have been calculated appropriately, particularly after revisions were made following the manufacturer's response to clarification questions. However, in some cases the ERG believes costs could have been calculated more appropriately.

Firstly, the cost of prednisone is included in the economic model for patients receiving RTX and patients receiving CYC. This reflects that prednisone was administered in the RAVE trial. However, the ERGs clinical advisors state that in the UK prednisolone rather than prednisone would be administered, and that 1mg of prednisone is equivalent to 1mg of prednisolone. The average cost per mg across each available oral formulation of prednisone is £0.43, compared to £0.02 for prednisolone.[63] Because prednisolone (or prednisone) is given alongside both RTX and CYC the relative effect of applying this cost change in the manufacturer's economic model is relatively small. However, because more courses of treatment are assumed to be given in the RTX arm of the model, the ICER is reduced.

Secondly, for RTX, CYC (oral and IV), AZA and trimethoprim costs were based upon recommended or licensed doses rather than the doses observed in the RAVE trial (see MS p.154-155),[8] whereas for methylprednisolone and prednisone costs were based upon the average amount received in the RAVE

Superseded – see erratum

study, with treatment groups combined (that is, it was assumed that the dose received was equal in patients treated with RTX and patients treated with CYC). It is logical to base methylprednisolone and prednisone doses on what was received in RAVE, since these treatments were tapered in patients who achieved remission and therefore part of the impact of a successful treatment is to reduce the use of these. Given this, it appears strange that a weighted average dose of methylprednisolone and prednisone was applied to both arms of the economic model – it would appear more appropriate to base the dose in each modelled treatment arm on the data from that treatment arm observed in the RAVE trial. However, Stone *et al.* state that while the prednisone and methylprednisolone doses appeared to be slightly lower in the RTX group in the RAVE trial, post-randomisation use was conditional on pre-randomisation use and thus the difference could not be clearly attributed to treatment difference.[78] Therefore the manufacturer's approach seems reasonable, and potentially conservative.

Thirdly, the manufacturer's approach to estimating the drug costs associated with RTX and CYC may be biased in favour of the RTX group. It is assumed that patients receive a full 6-month course of RTX and CYC treatment – that is four 375mg/m^2 infusions of RTX, 2mg/kg/day of oral CYC and 10 infusions of 15mg/kg of IV CYC. Although data provided by the manufacturer on the RAVE trial shows that 95% of patients in the RTX group and 99% of patients in the CYC group received at least 75% of the planned dose, the mean cumulative doses appear to tell a different story. The mean cumulative dose of RTX during the original treatment period was 1,478.52mg/m^2 (standard deviation 123.63), very close to the 1,500mg/m^2 that would have been received had the full dose been received. On the other hand, the mean cumulative dose of CYC was 15,234.02mg (standard deviation 7,199.724), which, given the mean weight in the CYC group in RAVE of 87.88kg, is substantially below the full oral dose (32,098mg). It therefore appears that while it is reasonable to expect that patients treated with RTX will receive very close to the total dose recommended, for CYC this may not be true. The cost of CYC is relatively low, but assuming that the dose received is approximately half that suggested by a "full" dose increases the ICER appreciably, particularly if it is also assumed that the number of administrations of IV CYC is also approximately halved – however clinical advice received by the ERG suggests that if IV CYC is given it is more likely that a full 6-month course would be administered, though this may not always be the case, especially if patients achieve remission before the 6-month timepoint. This is discussed further in Section 6.1.4.

Fourthly, patients taking oral or IV CYC are assumed to receive 400mg of trimethoprim as pneumocystis jiroveci prophylaxis based upon BSR guidelines (see MS p.152). It is assumed that this prophylaxis is not given to patients treated with RTX. However, such treatment is recommended in the draft SmPC for RTX, as stated on p.19 of the MS.[8] The cost of trimethoprim is very low (£21.38

Superseded – see erratum

per cycle) and hence the inclusion of this cost in the RTX arm of the model will have a minimal impact. However, the ICER will increase marginally.

In addition, the weight and BSA assumed in the manufacturer's model impacts upon the cost-effectiveness results. The mean BSA observed in the RAVE trial was 2.00m^2 and the mean weight was 85.1kg (according to RAVE data on file provided by the manufacturer, though this was stated to be 87kg in the MS (see MS p.138)).[8,21] The manufacturer claims that these are unrealistically high and instead use a BSA of 1.79m^2 and a weight of 67.2kg in the economic model. The BSA figure is from a paper that attempts to estimate realistic BSAs for cancer patients in the UK (Sacco *et al.*).[56] However, not only are cancer patients likely to be different to vasculitis patients, the median age of the 3,613 patients included in the Sacco *et al.* study was 61 (the mean age is not given), compared to the mean age of 52.8 years in the RAVE trial. In addition, 40.7% of the patients included in the Sacco *et al.* study were male, compared to 50.3% in the RAVE trial. Sacco *et al.* found that males had higher BSAs than females, and that there was a negative correlation between BSA and age. Hence it seems likely that the BSA assumed for patients included in the manufacturer's model could confidently be seen as an underestimate if the model was addressing a cancer population. Given that the model is addressing an ANCA-associated vasculitis population the suitability of the BSA assumption becomes even more uncertain, but even more likely to represent an underestimate. The source for 67.2kg weight figure used in the manufacturer's model is not clear – it appears likely to have been taken from the Sacco *et al.* study, but a figure for average weight is not stated in that paper. Given that the assumed BSA appears to represent an underestimate, it is likely that the assumed weight is also underestimated. Increasing the weight and BSA included in the model to the averages observed in the RAVE trial leads to an appreciable increase in the ICER.

The ERGs clinical advisors were concerned about the manufacturer's assumption that 72% of patients treated with CYC received IV treatment, with 28% receiving oral treatment, based upon unreferenced market research (see MS p.161).[8] Given the lower adverse event risk associated with IV CYC, and the lower cumulative dose allowing additional courses of treatment, the ERGs clinical advisors felt that IV CYC should be considered the primary comparator.

Finally, the ERG is concerned that wastage has not been included in the manufacturer's base case analysis, and that when wastage is included in sensitivity analysis it is not incorporated accurately, particularly because the required dose is dependent upon the assumed BSA. With an assumed BSA of 1.79m^2 the required RTX dose is 671mg per administration, thus one 500mg vial and two 100mg vials are sufficient. This is associated with a cost per course of £4,889.64, which is only marginally higher than the cost based upon no wastage of £4,689.78 – hence in this case including wastage does not have a substantial impact upon the ICER. However, BSAs of 1.87m^2 and above require RTX doses of more than 700mg per administration, thus one 500mg vial and three 100mg vials are required. Given

Superseded – see erratum

that the BSA of 1.79m^2 appears to be an underestimate and that the mean BSA observed in RAVE was 2.00m^2, this scenario seems possible. In this case, the cost per cycle of RTX increases to £5,588.16 which has the impact of almost doubling the manufacturer's base case ICER (with all other parameters remaining as assumed by the manufacturer). Ideally, information on the distribution of BSAs (and on patient weights in order to inform CYC IV dose wastage) from the RAVE trial would be used to calculate wastage. Calculating wastage based upon mean weight and mean BSA is likely to be inaccurate, since some patients may "only just" require an additional vial, whereas others may use almost all of their final vial. Ideally, the manufacturer would present mean cost information for all weight- or BSA-related treatments based upon the actual number of vials required for each patient in the RAVE trial, including wastage. It should be noted that while IV CYC wastage is relevant it is of much lesser importance due to the extremely low cost associated with vials of CYC. In addition, the number of CYC vials required by a patient weighing 67kg and a patient weighing 85kg is the same, and thus the manufacturer's wastage analysis for CYC is not affected by weight assumptions.

Administration costs

Administration costs were included for RTX and IV CYC. The ERG is concerned that administration costs were based upon the cost estimated by Barton *et al.*[64] for an administration of infliximab. The original reference presented a cost for a price year of 2000, and the manufacturer uprated this to 2012 using the Hospital and Community Health Services Pay and Prices Index presented by Curtis (2012).[68] Barton *et al.* do not provide a source or a method for their administration cost estimate.

The MS states that RTX should be given in hospital as a day case procedure (see MS p.28).[8] The average weighted cost of all day case activity reported in the NHS Reference Costs in 2009-10 (the source of costs used in the manufacturer's model) was £673.20, although the ERG accepts that none of the HRGs for which data are presented appear to be relevant for the infusion of RTX. The unit cost of delivering a subsequent element of a chemotherapy cycle (which may include the administration of RTX in cancer patients) was £284.45.[65] It is notable that in a recent NICE Single Technology Appraisal of golimumab in rheumatoid arthritis sensitivity analysis on the administration cost of RTX included a cost of £284.73 per administration, though the source of this value is unclear.[79] The ERGs clinical advisors suggest that assigning the same administration cost to RTX and CYC may cause some bias in favour of RTX, since RTX typically involves a longer infusion time. The impact of relaxing this assumption is investigated in Section 6.1.4.

The ERG note that the manufacturer assumes that methylprednisolone is given at the same visit as the first infusion of RTX (the same assumption is made for IV CYC), and no additional administration costs are included. This is said to be a conservative assumption because no administration cost is associated with methylprednisolone even when CYC is given orally. The draft SmPC for RTX states

Superseded – see erratum

that 1,000mg of methylprednisolone should be administered for 1-3 days prior to the first dose of RTX, with the final dose given on the same day as the first dose of RTX.[80] Figures included in the economic model suggest that the mean methylprednisolone dose was 1,627mg in the RTX group and 1,693.9mg in the CYC group. The ERG believes that this indicates that at least 1 additional administration cost should be applied for methylprednisolone as part of RTX and CYC treatment.

Monitoring costs

Monitoring costs were included in the manufacturer's model for oral CYC and AZA. No additional monitoring costs were included for RTX or IV CYC because the manufacturer assumes that any monitoring is undertaken during the administration of these drugs.

The manufacturer assumes that during oral CYC treatment patients receive 9 blood tests and 9 LFTs in each 6 month cycle that treatment is received, based upon Lapraik *et al.* (2007).[9] In fact, Lapraik *et al.* state that full blood tests should be given weekly for 1-month when oral CYC is being given. Subsequently full blood tests should be undertaken every two weeks for 2 months, and thereafter tests should occur once per month. The ERG suggest that this would be equivalent to approximately 11 full blood tests per 6 month cycle. In addition Lapraik *et al.* do not state that LFTs should be undertaken in these patients, instead stating that renal function tests should be conducted alongside the full blood tests. Therefore it is not clear why the manufacturer has made the assumptions that they have made regarding the monitoring of oral CYC treatment.

The manufacturer assumes that during AZA treatment patients receive 9 blood tests and 9 LFTs in each 6 month cycle that treatment is received, based upon Chakravarty *et al.* (2008 and 2009).[69] In fact, Chakravarty *et al.* state that full blood tests and LFTs should be given weekly for 6 weeks when AZA is being given. Subsequently full blood tests and LFTs should be undertaken every two weeks for 6 weeks, and thereafter tests should occur once per month. The ERG suggest that this would be equivalent to approximately 12 full blood tests and LFTs per 6 month cycle. Chakravarty *et al.* also state that creatinine and urea and electrolytes should be tested once every 6 months. Therefore it is not clear why the manufacturer has made the assumptions that they have made regarding the monitoring of AZA treatment.

The manufacturer assumes that the cost of a full blood test is £3, based upon NHS Reference Costs (currency code DAP823),[65] and that the cost of LFTs is £9, based upon uprating to current prices the cost of £6.19 reported by Barton *et al.* (2004).[64] Barton *et al.* estimated their cost based upon data provided by the University Hospital Birmingham NHS Trust in 2000. NHS Reference Costs report a cost of £1.29 for biochemistry pathology services (currency code DAP841[65]) which may be appropriate for LFT and renal function tests.

Superseded – see erratum

The assumption that monitoring is assumed to take place during RTX and IV CYC administrations is debatable, given that clinicians are likely to wish to see the results of monitoring tests before continuing with treatment. Hence the ERG suggest that a more appropriate assumption may be that patients who take RTX or IV CYC receive the same number of blood tests and renal function tests as the number of administrations of their treatment.

The ERG acknowledge that the number and type of monitoring tests received will likely vary depending upon individual patient characteristics. As well as LFT and renal function tests, CRP, plasma viscosity, ESR, ANCA, urine protein and a range of other measures are often taken routinely. The ERG has attempted to estimate a reasonable number of tests for each treatment per cycle, but it is important to note that due to their low cost these are highly unlikely to have a major impact on cost-effectiveness results.

Health state costs

Health state costs form by far the largest portion of total costs generated by the manufacturer's economic model – 93% for the CYC group in the manufacturer's base case "all patients" analysis, and 89% for the RTX group. Therefore these costs have an exceptionally important effect on the cost-effectiveness results. It is the opinion of the ERG and our clinical advisors that these costs have been substantially over-estimated, creating a significant bias in favour of RTX.

It is assumed that patients in the "non-remission" health state attend consultant-led outpatient appointments once every 1.5 weeks, but because of multi-organ involvement it is essentially assumed that patients in this state have 3.1 outpatient appointments every 1.5 weeks (53.8 appointments per 6 month cycle). It is assumed that patients in the "remission" health state attend appointments once every 3 months, but because of multi-organ involvement it is essentially assumed that patients in this state have 3.1 outpatient appointments every 3 months (6.2 appointments per 6 month cycle). In the "uncontrolled disease" health state it is assumed that patients attend one "specialist palliative care" outpatient appointment (NHS reference cost SD04A[65]) every 1.5 weeks for the remainder of their life (17.4 appointments per 6 month cycle) (see MS p.158).[8] This cost is of particular importance given that in the manufacturer's model patients in the "standard care" CYC group spend 70.7% of their discounted mean life expectancy in the "uncontrolled disease" health, compared to 63.2% in the RTX group.

Clinical opinion received by the ERG suggests that for patients in the "remission" health state patients could be expected to have outpatient appointments once every 3 months for the first year. This would then likely reduce to one appointment every 6-12 months. For patients in the "non-remission" health state the ERG's clinical experts suggest that one outpatient appointment every 2-3 weeks is realistic.

Superseded – see erratum

For patients in the "uncontrolled disease" health state, which the ERG liken more to a low-grade "grumbling" disease health state, our clinical advisors suggest that patients may be seen once per month for 3-4 months, followed by appointments once every 3 months subsequently. In this state the ERG's clinical advisors suggest that patients are likely to be receiving treatment to maintain some control over their disease. Possible treatments are wide-ranging, but an assumption that these patients continue to take 2mg/kg of AZA per day appears to represent a reasonable proxy for the fact that these patients will be receiving treatment. Even assuming that patients in the "uncontrolled disease" health state continue to have outpatient appointments once per month for the remainder of their lives as well as AZA treatment greatly reduces the health state cost. This has the impact of substantially increasing the manufacturer's base case ICER (holding all other parameter assumptions constant).

In addition to the outpatient costs assumed in the manufacturer's model, it is assumed that patients with pulmonary symptoms (53% of patients based upon baseline characteristics in the RAVE trial) receive chest X-rays or CT scans once every 1.5 months while in the "non-remission" and "remission" health states. The cost for these was assumed to be £29.08, based upon Barton *et al.* (2004) and uplifted to 2012 prices. However, according to NHS Reference Costs the cost of diagnostic imaging in a consultant led follow-up attendance, non-admitted face-to-face attendance is £18.56 (service code 812), while the cost of a CT scan of one area with no contrast in the outpatient setting is £100.41 (currency code RA08Z).[65] The ERG considers that these are more relevant unit costs for this appraisal.

5.3.4 Technical issues relating to the implementation of the model

The ERG identified a number of technical issues within the manufacturer's model. The majority of these are minor and are unlikely to have a significant impact on the ICER. However, issues around the random number sampling and the relatively low number of probabilistic iterations included in the PSA mean that the manufacturer's PSA results should be considered unreliable.

Mortality estimates

In the manufacturer's model, the probability of death associated with the "uncontrolled disease" health state becomes 1.0 prematurely (in cycle 78, with patients aged 91). In their response to clarification questions the manufacturer stated that this was because at this point the probability of death in the general population (0.1676) multiplied by the SMR used for patients in the "uncontrolled disease" health state led to a probability of greater than 1.0.[20] However, the SMR quoted in the manufacturer's response to clarification questions is 6.6, whereas the SMR used in the model for patients in the "uncontrolled disease" health state is 5.28. 5.28 multiplied by 0.1676 is less than 1, hence the manufacturer's model is incorrect. However, this has very little impact on the ICER.

Superseded – see erratum

Costs

In the manufacturer's model the cost of a blood test is given as £3.00, however the actual cost stated in the Reference Cost source used by the manufacturer is £3.06.[65] In addition this cost, and the cost of the liver function test, are not linked to the "monitoring cost" calculations made in the model, and thus are excluded from the PSA. The upper and lower quartile ranges for the DVT adverse event have also been inputted incorrectly – in the manufacturer's model these are £287.55 and £480.32, whereas in the Reference Cost source used by the manufacturer they are £274.16 and £471.76.

Parameter distributions

Beta distributions are used to characterise the uncertainty surrounding relapse rates. In the manufacturer's original submission it was assumed that standard errors were equal to 30% of the mean. Upon response to clarification questions this was amended, such that the α and β parameters were calculated using the mean estimated relapse rate and the number of patients included in each arm of the study. However, this does not represent the full uncertainty associated with these parameters because relapse rates were estimated based only upon patients who first achieved remission – not all patients included in the trial.

The manufacturer arbitrarily estimates that the standard errors associated with the remission risk ratios are 20% of their mean, for both CYC and RTX. Even for the remission rate associated with a second course of CYC, uncertainty is characterised based upon the combined uncertainty around the remission rate for a first course of CYC and the uncertainty around the risk ratio associated with a second course – rather than being directly based upon the uncertainty in the remission rate in patients with "recurrent disease" observed in the RAVE trial. The use of risk ratios allows the manufacturer to model multiple lines of CYC and RTX treatment. However, given the ERGs belief that it is not appropriate to model more than two courses of CYC (one for most patients with "recurrent disease"), and that it is not appropriate to model two courses of RTX treatment, the remission risk ratios become redundant – remission rate estimates can be taken directly from RAVE.

In addition, the beta distribution for one of the risk ratios incorporated within the manufacturer's model was capped such that values from the very top end of the distribution could not be sampled in order to avoid an Excel error. This is unlikely to have significantly affected the model results, and is no longer an issue given the recommendation made by the ERG that these risk ratios become redundant in the model.

The probabilistic values of liver function tests and imaging through chest X-ray or CT scan represent multiplications of the "blood test" cost, based upon their relative deterministic means. In their clarification response, the manufacturer explained that this was because no source could be found in

Superseded – see erratum

order to inform the uncertainty in the costs of these tests and scans. The ERG suggests that if Reference Cost sources had been used for these rather than uprated figures from the Barton *et al.* study uncertainty could have been taken into account more appropriately.

The ERG also note that for all cost parameters the manufacturer used log normal distributions and that the distributions of these were estimated incorrectly due to inappropriate estimations of the standard error. The ERG suggests it is reasonable to assume that costs informed by reference costs are normally distributed, and the uncertainty in these can be characterised based upon the number of data submissions that informed the unit cost estimate and the inter-quartile range.

The manufacturer used beta distributions to characterise the uncertainty around health state utility scores. Standard errors for these distributions were estimated by dividing standard deviations by the square root of the sample size of patients that informed the utility estimate in each state. The health state utility scores were not correlated and the "non-remission" and "remission" distributions overlap. However this is marginal – the ERG find that in 9 iterations of 5,000 the "non-remission" health state utility score was higher than the "remission" state utility score. The ERG believes that this is an adequate representation of the uncertainty in the utility associated with these health states. However, the probabilistic value of the "uncontrolled disease" health state is programmed to always equal the utility in the "non-remission" health state, minus the difference between utility in the "remission" health state and the "non-remission" health state. The ERG suggest that this does not appropriately reflect the uncertainty in the "uncontrolled disease" health state, since it is uncertain whether the difference between the "remission" and "non-remission" states is the same as the difference between the "non-remission" and "uncontrolled disease" states.

The manufacturer used beta distributions to characterise the uncertainty around adverse event utility scores. Because of the small part played in the model by adverse events, these are of relatively low importance; of more are the inconsistent sources and potentially inappropriate durations assumed for these utilities (as discussed in Section 5.2.4). For anaemia, the standard error of the utility score was arbitrarily assumed to equal 30% of the mean, whereas for the other AEs standard errors were derived from the data source.

Finally, the ERG notes that uncertainty was not characterised around the following parameters:

- Standardised mortality ratios
- Drug, administration and monitoring costs
- Proportion of patients visiting consultants of different specialty and receiving tests
- Frequency of outpatient appointments by health state
- Frequency of X-rays and scans for people with pulmonary involvement

Superseded – see erratum

The ERG suggests that the failure to characterise uncertainty around the frequency of outpatient appointment by health state is of particular importance for the probabilistic results of the economic model.

5.4 Summary of key issues and uncertainties within the manufacturer's model

The MS presented a *de novo* model to estimate the incremental cost-effectiveness of a treatment sequence beginning with RTX compared to a treatment sequence beginning with CYC in patients with severe ANCA-associated vasculitis. Based on the submitted version of the manufacturer's model, the treatment sequence including RTX is consistently expected to produce more QALYs than the sequence beginning with CYC. Additional costs are incurred in the RTX treatment sequence and the incremental cost-effectiveness of the RTX sequence compared to the CYC sequence is expected to be around £8,544 per QALY gained. Important subgroup analysis was undertaken on "treatment naïve" patients, patients with "recurrent disease" and patients intolerant to CYC treatment. The manufacturer estimated ICERs of £55,175 per QALY gained and £43,003 per QALY gained for the "treatment naïve" and "recurrent disease" subgroups respectively, and estimated that a RTX treatment sequence would dominate a best supportive care sequence for patients who are intolerant to CYC.

The ERG notes that all of the manufacturer's cost-effectiveness results are highly sensitive to assumptions made about the number of courses of RTX and CYC received in each treatment pathway. Based upon clinical expert advice, the ERG believes that more appropriate treatment sequences exist and that these have not been modelled by the manufacturer; consequently the results presented by the manufacturer should be approached with considerable caution. The manufacturer's model has sufficient flexibility built in such that alternative sequences can be assessed with relatively little additional programming required – the results of these analyses (combined with other alterations to the manufacturer's model) are presented in Section 6.1.3.

In addition, the manufacturer's cost-effectiveness results are highly sensitive to assumptions made about the costs and quality of life associated with the "uncontrolled disease" health state. The ERG believes that ideally the manufacturer's model would have included additional lines of therapy – with treatments such as MMF, leflunomide, AZA, MTX and others. Expert clinical advice received by the ERG suggests that these treatments, or combinations of them, are likely to play a part in the lifetime treatment sequences received by patients with generalised, "severe" ANCA-associated vasculitis. Given that these treatments are available and are used, the ERG believe that it is unrealistic to assume that once patients have relapsed after receiving RTX and CYC no treatment options remain, leaving patients to live their lives with symptomatic disease – incurring high costs and a low quality of life.

Superseded – see erratum

The ERG believes that a better description of this health state would be low-grade "grumbling" disease, which is partially controlled through treatment. This state would involve a higher utility score than that assumed by the manufacturer for "uncontrolled disease", and although costs would be incurred these would not be as high as those assumed by the manufacturer – the ERG and our clinical advisors believe that the assumption that patients with "uncontrolled disease" have outpatient appointments to receive specialist palliative care every 1.5 weeks for the remainder of their lives represents a substantial over-estimate.

Overall, the ERG believes the following to represent the most important issues and uncertainties surrounding the manufacturer's submitted economic analysis:

- Several realistic treatment sequences were not modelled for the "all patients" analysis as well as the subgroup analyses.
- Inappropriate costs and (to a lesser extent) utilities were assumed for the "uncontrolled disease" health state (which could be more accurately described as "grumbling disease").
- An inappropriate assumption was made that all flares lead to immediate re-induction therapy – leading to an over-estimate of the relapse rate and unrealistically quick transition to the "uncontrolled disease" health state.
- Assumptions around the resource use costs associated with the "remission" and "non-remission" health states are questionable – the resource use assumed in the "non-remission" state in particular seems to be considerably over-estimated.
- Inappropriate assumptions were made around weight, BSA and wastage. Weight and BSA seem to be underestimated, and wastage is not included in the base case analyses.
- The manufacturer assumed that the glucocorticoid prednisone would be given alongside CYC or RTX, rather than prednisolone. In a UK context, this is inappropriate.
- The manufacturer considerably over-estimated the amount of oral CYC used in a typical treatment course.
- Several important parameters were not included in the PSA conducted by the manufacturer.

The potential impacts of these issues are explored in the next section.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

This section presents additional work undertaken by the ERG to explore the impact of the key issues identified through the critical appraisal of the manufacturer's model. Section 6.1 details the amendments made to correct the manufacturer's model and the additional economic analyses undertaken by the ERG. Section 6.2 presents a discussion of the key uncertainties surrounding the cost-effectiveness of RTX.

It is important to note that some of the issues highlighted within the critical appraisal (Section 5) relate to apparent mistakes in the programming of the model, whilst others concern matters of judgement with respect to the structure of the model and the evidence used to inform its parameters. The ERG-corrected ICERs presented in Section 6.1.1 reflect the former, applied to the treatment sequences applied by the manufacturer. The ICERs presented in Section 6.1.2 reflect the view of the ERG and our clinical advisors with respect to the values of parameters used within the model, applied to the treatment sequences outlined by the manufacturer. The ICERs presented in Section 6.1.3 reflect a full incremental analysis of the treatment sequences outlined as appropriate in Section 5.3.2, combined with the view of the ERG and our clinical advisors with respect to the values of parameters used within the model. In Section 6.1.4 a set of final scenario analyses are presented, in line with comments made by the ERG's clinical advisors.

6.1 Additional work on cost-effectiveness undertaken by the ERG

6.1.1 Methods used within the ERG's additional analyses – correcting mistakes

List of model amendments made by the ERG

The ERG made the following amendments to the manufacturer's model to rectify apparent programming (technical) errors.

(i) **Mortality risk** – Mortality risks for patients aged 91 and over in the "uncontrolled disease" health state were amended.

(ii) **Cyclophosphamide cost –** The cost of oral CYC was altered in order to match the mean dose received in the RAVE trial, rather than a full 6-month dose. In this analysis it is assumed that 10 courses of IV CYC are received – this assumption is tested further in Section 6.1.4.

(iii) **Prednisolone cost –** The model was altered to include the cost of prednisolone, rather than the cost of prednisone.

(iv) **Pneumonia utility –** The utility score applied to pneumonia adverse events was adjusted to reflect that the source referred to a utility score experienced for 34 days, rather than 6 months.

(v) **Trimethoprim cost –** The cost of trimethoprim was added to the RTX group in line with the draft SmPC.

(vi) **Blood test cost –** The cost of blood tests was amended to match that quoted in the cost source.

(vii) **Deep Vein Thrombosis cost –** The upper and lower limits of this cost were amended to match those quoted in the cost source.

(viii) **Relapse rate numbers at risk –** These were altered to match the numbers who achieved remission in each group.

(ix) **Distributions for cost parameters –** These were altered to normal distributions.

(x) **Distributions for standardised mortality rates and outpatient appointments –** The model was altered such that these parameters were included in the PSA.

The stepwise cumulative impact on the deterministic ICER of each of these amendments is presented in Table 16, for the "all patients" population.

Table 16: Correcting mistakes in the manufacturer's economic model

Alteration Number	Alteration	Inc. costs	Inc. QALYs	ICER
0	Base Case	£1391.03	0.1628	**£8,543.69**
1	Mortality rates calculation	£1391.03	0.1628	**£8,543.69**
2	CYC oral cost	£1401.87	0.1628	**£8,610.27**
3	Prednisolone cost	£959.07	0.1628	**£5,890.61**
4	Pneumonia AE utility	£959.07	0.1642	**£5,841.13**
5	Trimethropim cost	£987.78	0.1642	**£6,015.97**
6	Blood test cost and link this and LFT cost to model calculations	£986.13	0.1642	**£6,005.92**
7	DVT upper and lower limit costs	£986.13	0.1642	**£6,005.92**
8	Relapse rate beta distribution "n's"	£986.13	0.1642	**£6,005.92**
9	Normal distributions for cost parameters	£986.13	0.1642	**£6,005.92**
10	Make SMRs, consultant visits probabilistic	£986.13	0.1642	**£6,005.92**

Table 16 demonstrates that the corrections made to the manufacturer's model generally led to reductions in the ICER – particularly the inclusion of costs for prednisolone rather than prednisone. This is also true for the "treatment naïve" and "recurrent disease" subgroup analyses, with base case ICERs falling from £55,175 per QALY gained and £43,003 per QALY gained (based upon the manufacturer's base case analyses) to £47,496 per QALY gained and £37,970 per QALY gained respectively. The probabilistic ICER for the "all patients" analysis fell to £4,437 per QALY gained

(run with five-thousand samples) – notably lower than the deterministic analysis. This is primarily due to the wide distributions placed around the relapse rate parameters – when these are removed from the probabilistic analysis the probabilistic ICER is very similar to the deterministic ICER (£6,051 per QALY gained).

6.1.2 Methods used within the ERG's additional analyses – altering parameter values

List of model amendments made by the ERG

(i) **Body surface area (BSA) and weight** – BSA and weight used in the model were increased to 1.83m^2 and 70.51kg respectively. These values were estimated by fitting a linear regression to the raw data from the Sacco *et al.*[56] study in order to accurately account for the mean age and proportion of males in the RAVE trial.

(ii) **Maintenance therapy in the rituximab group –** Patients induced into remission by RTX therapy are assumed to receive AZA maintenance therapy, using the same dose as received in the CYC group.

(iii) **Relapse rates –** Relapse rates were re-estimated based upon data on severe flares in the CYC group of the RAVE trial. This reflects an assumption that only severe flares will lead to renewed induction treatment. Given the ERG's assumption that patients in the RTX group receive AZA maintenance treatment, the same relapse rate is applied to patients in the RTX group and patients in the CYC group. The relapse rate is assumed to be identical after subsequent lines of therapy.

(iv) **Costs in "uncontrolled disease" state –** To reflect clinical advice received by the ERG patients in the "uncontrolled disease" health state are assumed to have a degree of disease control (this is assumed to be low-grade, "grumbling" disease), and are treated with AZA. As such these patients also incur monitoring costs. One outpatient appointment with a consultant is assumed to occur every 2 months.

(v) **Utility in "uncontrolled disease" state –** To reflect the ERG's assumption that patients in this health state have a degree of disease control it is assumed that the difference in the utility score experienced in this state compared to the "non-remission" state is not the same as the difference in utility score between the "remission" state and the "non-remission" state. Instead it is arbitrarily assumed that the difference in the utilities experienced in the "uncontrolled disease" and the "non-remission" health states is half the difference between the utilities experienced in the "remission" and "non-remission" health states. In the base case, this is equivalent to utility scores of 0.84, 0.75 and 0.71 in the "remission", "non-remission" and "uncontrolled disease" health states respectively.

(vi) **Numbers and costs of blood tests, liver function tests (LFTs) and renal function tests –** The number of these tests were re-estimated to reflect those recommended in the Lapraik *et al.* and Mukhtyar *et al.* guidelines. The ERG acknowledges that the precise

number and type of monitoring tests received will likely vary depending upon patient characteristics. The ERG has attempted to estimate a reasonable frequency of these tests, as shown in Table 17. Due to their low costs, these are highly unlikely to have important impacts upon the cost-effectiveness results.

Table 17: Tests and costs

Treatment	Blood tests		Liver function tests		Renal function tests	
	Number	Cost	Number	Cost	Number	Cost
CYC (oral)	11	£3.06	0	£1.29	11	£1.29
AZA	12	£3.06	12	£1.29	1	£1.29

(vii) **Methylprednisolone administration** – While the ERG accept that the final methylprednisolone dose may be given on the same day as the first RTX or IV CYC dose, it is noted that in the RAVE trial 1-3 pre-treatment doses were administered. Therefore the cost of an administration of methylprednisolone was added to the CYC and RTX groups. It was assumed that the cost of this was equal to the cost of administering CYC or RTX.

(viii) **Cost of X-rays and CT scans** – In the manufacturer's model one cost was used (£29) to estimate the cost of an X-ray or a CT scan, based upon uplifting a figure taken from Barton *et al.* In the ERG's amended analyses these are instead based upon the cost of relevant NHS Reference Costs (£18.56 for an X-ray and £100 for a CT scan).[65] Based upon clinical advice it was assumed that 80% of scans received in the modelled population would be X-rays, with 20% being CT scans.

(ix) **Wastage** – In the ERG's base case analyses the costs of wastage for RTX and IV CYC are included. These are estimated based upon the weight and BSA specified in (i), above.

(x) **Number of outpatient appointments in health states** – Clinical advice received by the ERG suggests that the number of outpatient appointments assumed in the manufacturer's economic model are considerable over-estimates – particularly in the "uncontrolled disease" health state. These have been substantially altered, as shown in Table 18.

Table 18: Outpatient appointments

Health state	No. of outpatient appointments per 6 months – manufacturer's assumptions	No. of outpatient appointments per 6 months – ERG's assumptions
Remission	6.19	2.00
Non-remission	53.82	13.03
Uncontrolled disease	17.38	3.00

The stepwise impact of the alterations made to the manufacturer's model by the ERG is presented in Table 19. The results refer to the "all patients" base case analysis.

Table 19: Altering parameters in the manufacturer's economic model

Alteration Number	Alteration	Inc. costs	Inc. QALYs	ICER
10	From Table 16	£986.13	0.1642	**£6,005.92**
11	Increases BSA/Weight	£1,105.33	0.1642	**£6,731.89**
12	RTX maintenance with AZA	£1,729.82	0.1642	**£10,535.29**
13	Amended relapse rates	-£1,773.28	0.2700	**RTX dominates**
14	Amended costs in uncontrolled disease	-£3,002.36	0.2700	**RTX dominates**
15	Amended utility in uncontrolled disease	-£3,002.36	0.2165	**RTX dominates**
16	Amended blood test, LFT and renal function test numbers and costs	-£577.84	0.2165	**RTX dominates**
17	Include methylprednisone administration	-£527.74	0.2165	**RTX dominates**
18	Amended cost of X-rays and CT scans	-£492.58	0.2165	**RTX dominates**
19	Include wastage	-£388.88	0.2165	**RTX dominates**
20	Amended number of outpatient appointments in health states	£5,704.14	0.2165	**£26,346.53**

Table 19 demonstrates that the alterations to parameter values made by the ERG have important effects upon the ICER. While allocating AZA maintenance therapy costs to the RTX group (alteration 12) increases the ICER, assuming the same relapse rates for RTX and CYC groups reduces the incremental costs associated with the RTX group to the extent that the RTX sequence dominates the CYC sequence (RTX is estimated to be both more effective and less costly). However, this result is substantially altered when assumptions around the number of outpatient appointments experienced in each health state are changed. This reflects that the main benefit associated with the RTX treatment sequence is that an additional line of treatment exists before patients reach the "uncontrolled disease" health state, and in the manufacturer's model the "uncontrolled disease" health state is very costly due to the assumed 17.38 outpatient appointments per 6 month period. Substantially reducing the cost of this health state causes the benefits associated with the RTX sequence of treatments to be considerably diminished.

6.1.3 Methods used within the ERG's additional analyses – treatment sequences

In Section 5.3.2 several relevant treatment sequences were outlined that should be analysed in order to determine i) Does the inclusion of RTX in the treatment sequence increase health benefits compared to the current treatment sequence? and ii) If so, where is the most cost-effective place in the pathway to position RTX?

In this section an incremental analyses of each relevant treatment sequence for each patient subgroup is presented, using an adapted version of the manufacturer's model including the amendments described in Sections 6.1.1 and 6.1.2.

All Patients

Table 20: ERG preferred analyses cost-effectiveness results – All patients

Strategy	Total Cost	Total QALYs	Inc. QALYs	Inc. Costs	ICER
CYC → CYC → BSC	£18,926.57	8.5810	-	-	-
CYC → CYC → RTX → BSC	£22,820.93	8.9035	0.32	3,894.36	**£12,075.42**
CYC → RTX → CYC → BSC	£23,176.00	8.9086	0.0051	355.07	**£69,709.63**
RTX → CYC → CYC → BSC	£23,755.25	8.9131	0.0045	579.25	**£127,456.12**

Table 20 demonstrates that for the "all patients" analysis, adding RTX to the treatment sequence after two courses of CYC treatment is associated with an ICER of £12,075.42 per QALY gained. Moving RTX forwards in the treatment sequence is associated with additional costs and marginal QALY gains. The ICER associated with administering RTX after one course of CYC is £69,709.63 per QALY gained compared to administering RTX after two courses of CYC, and the ICER of administering RTX as the first line of treatment is associated with an ICER of £127,456.12 per QALY gained compared to administering RTX at second line. Cost effectiveness acceptability curves (CEACs) and the cost effectiveness acceptability frontier (CEAF) associated with this analysis are presented in Figures 3 and 4 respectively. At a willingness-to-pay cost-effectiveness threshold of £30,000 per QALY gained, the probability that administering RTX after two courses of CYC is a cost-effective strategy in the "all patients" analysis is 58.3%. The probability that not including RTX in the treatment sequence represents a cost-effective strategy is 11.7%.

Figure 3: Cost effectiveness acceptability curves – All patients analysis

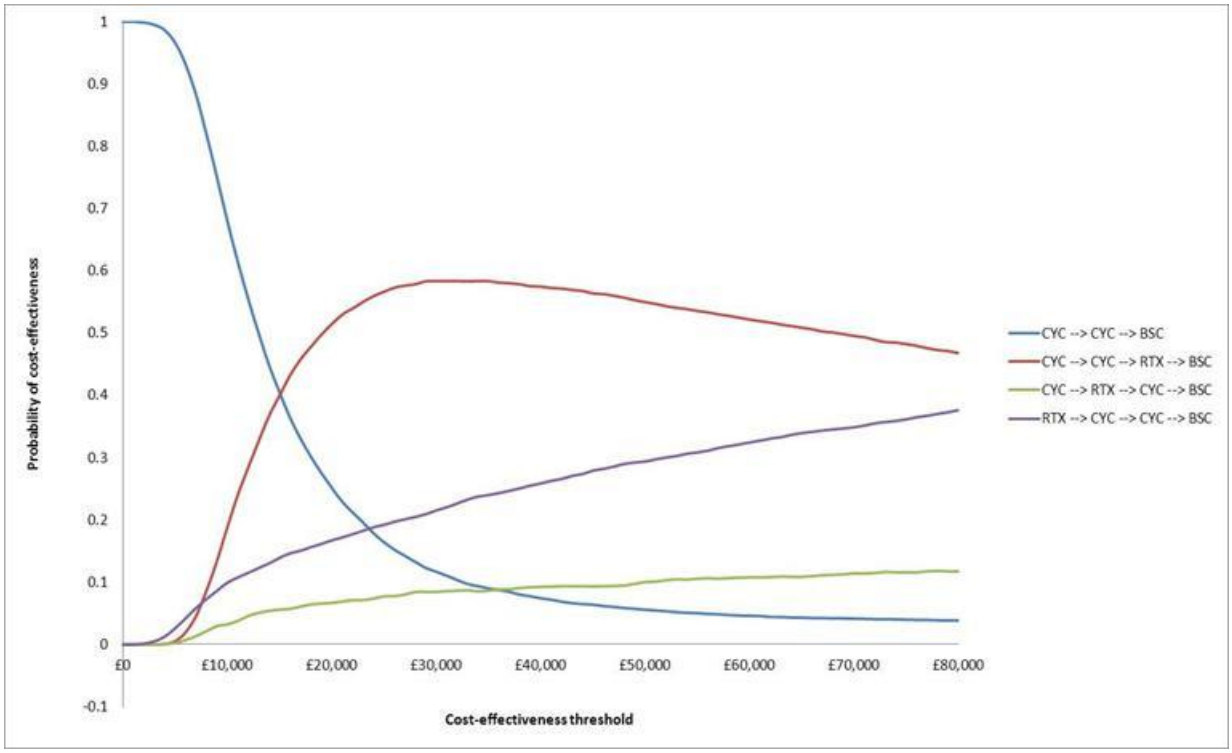
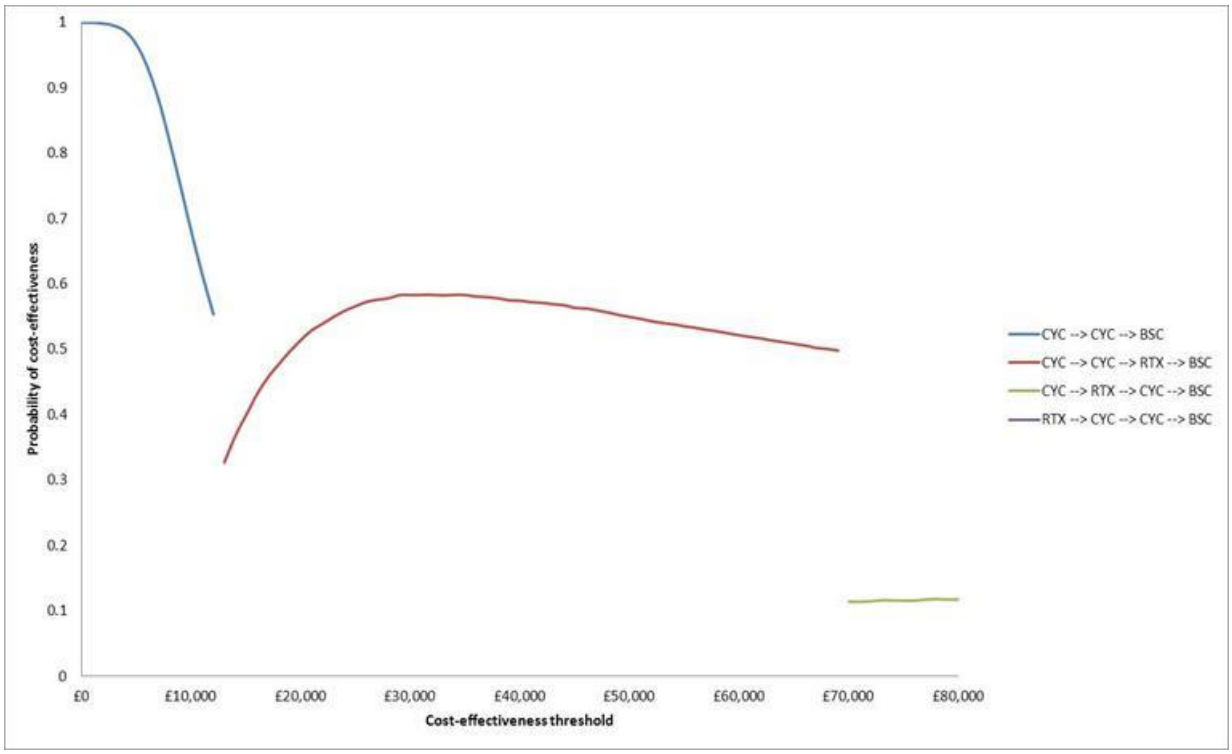


Figure 4: Cost effectiveness acceptability frontier – All patients analysis



Treatment naïve subgroup

Table 21: ERG preferred analyses cost-effectiveness results – Treatment naïve patients

Strategy	Total Cost	Total QALYs	Inc. QALYs	Inc. Costs	ICER
CYC → CYC → BSC	£18,645.81	8.6491	-	-	-
CYC → CYC → RTX → BSC	£22,429.08	8.9435	0.29	£3,783.27	**£12,850.76**
CYC → RTX → CYC → BSC	£22,793.54	8.9480	0.0045	£364.46	**£81,603.50**
RTX → CYC → CYC → BSC	£23,636.83	8.9507	0.0027	£843.29	**£317,037.96**

Table 21 demonstrates that for the "treatment naïve" analysis, adding RTX to the treatment sequence after two courses of CYC is associated with an ICER of £12,850.76 per QALY gained. Moving RTX forwards in the treatment sequence is associated with additional costs and marginal QALY gains. The ICER associated with administering RTX after one course of CYC is £81,603.50 per QALY gained compared to administering RTX after two courses of CYC, and the ICER of administering RTX as the first line of treatment is associated with an ICER of £317,037.96 per QALY gained compared to administering RTX at second line. CEACs and the CEAF associated with this analysis are presented in Figures 5 and 6 respectively. At a willingness-to-pay cost-effectiveness threshold of £30,000 per QALY gained, the probability that administering RTX after two courses of CYC is a cost-effective strategy in this analysis is 59.7%. The probability that not including RTX in the treatment sequence represents a cost-effective strategy is 13.9%.

Figure 5: Cost effectiveness acceptability curves – Treatment naïve subgroup

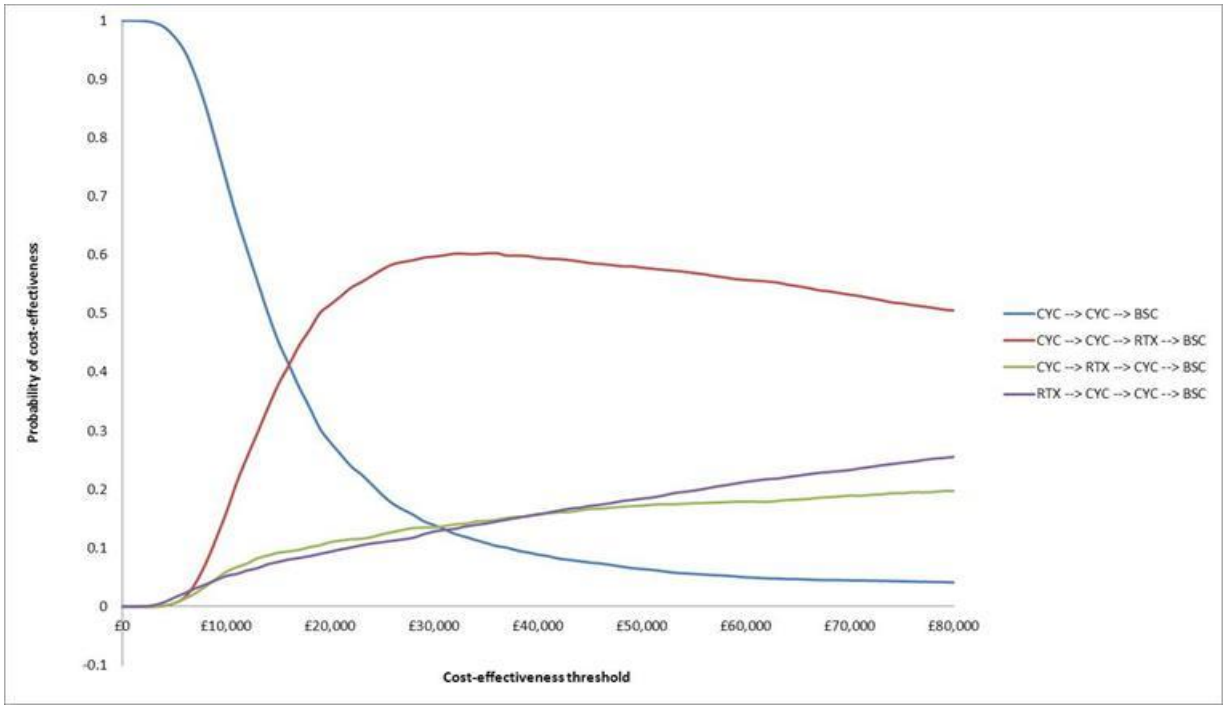
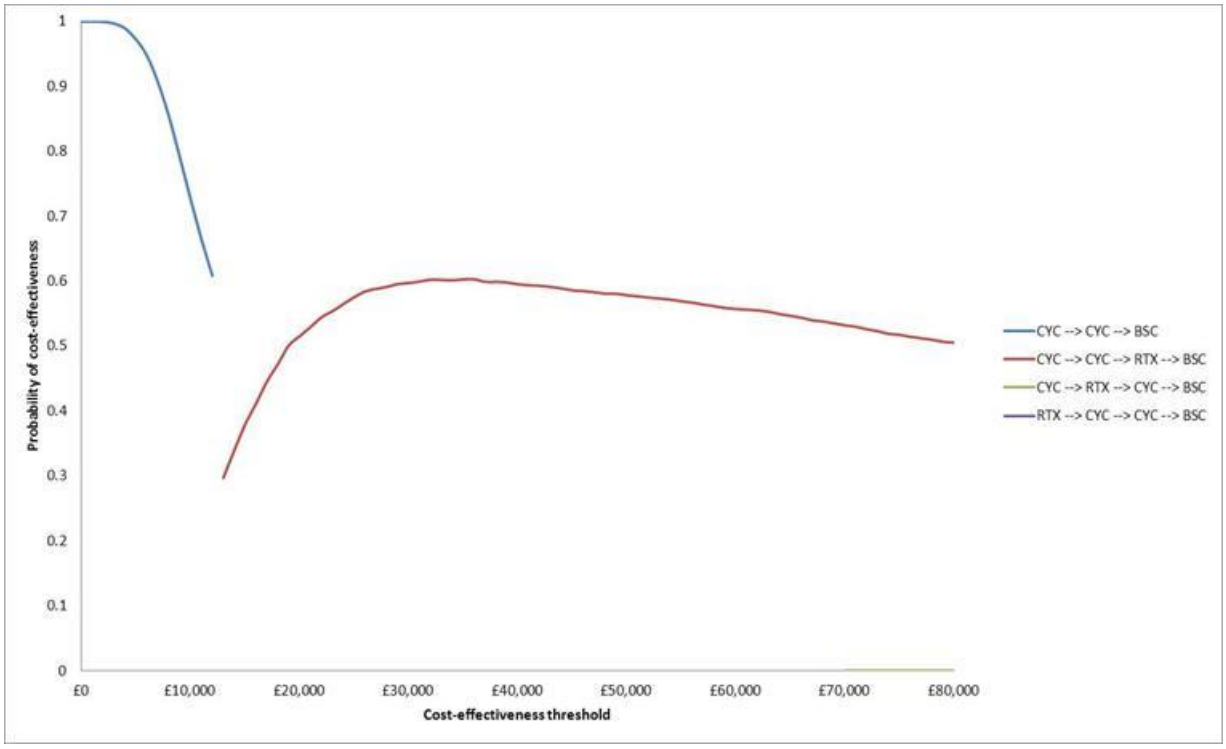


Figure 6: Cost effectiveness acceptability frontier – Treatment naïve subgroup



Recurrent disease subgroup – patients eligible for additional cyclophosphamide

For the "recurrent disease" subgroup it is relevant to consider two further sub-categories of patients: those who are eligible for further CYC treatment due to not yet surpassing the maximum recommended lifetime cumulative CYC dose of 20-30g; and those who are ineligible for further CYC treatment due to surpassing this maximum recommended cumulative dose. First results are presented for patients eligible for further CYC treatment.

Table 22: ERG preferred analyses cost-effectiveness results – Recurrent disease subgroup (eligible for additional cyclophosphamide treatment)

Strategy	Total Cost	Total QALYs	Inc. QALYs	Inc. Costs	ICER
CYC → BSC	£17,593.48	8.2548	-	-	-
CYC → RTX → BSC	£22,295.52	8.6773	0.4225	£4,702.04	**£11,129.22**
RTX → CYC → BSC	£22,620.65	8.6836	0.0063	£325.14	**£51,841.87**

Table 22 demonstrates that for the "recurrent disease" analysis (in which patients are eligible for additional CYC treatment), adding RTX to the treatment sequence after CYC is associated with an ICER of £11,129.22 per QALY gained. Moving RTX forwards in the treatment sequence is associated with additional costs and marginal QALY gains. The ICER associated with administering RTX as the first line of treatment is associated with an ICER of £51,841.87 per QALY gained compared to administering RTX at second line. CEACs and the CEAF associated with this analysis are presented in Figures 7 and 8 respectively. At a willingness-to-pay cost-effectiveness threshold of £30,000 per QALY gained, the probability that administering RTX after one course of CYC is a cost-effective strategy in this analysis is 51.3%. The probability that not including RTX in the treatment sequence represents a cost-effective strategy is 10.4%.

Figure 7: Cost effectiveness acceptability curves – Recurrent disease subgroup (eligible for cyclophosphamide)

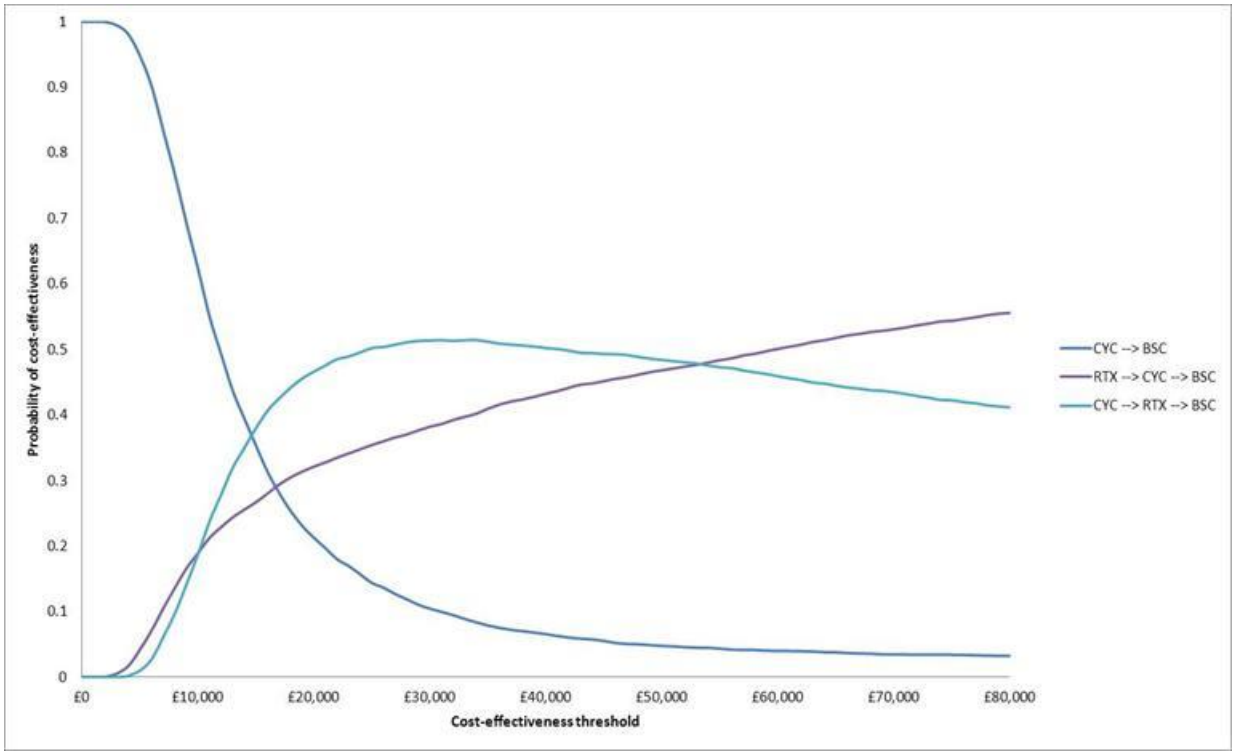
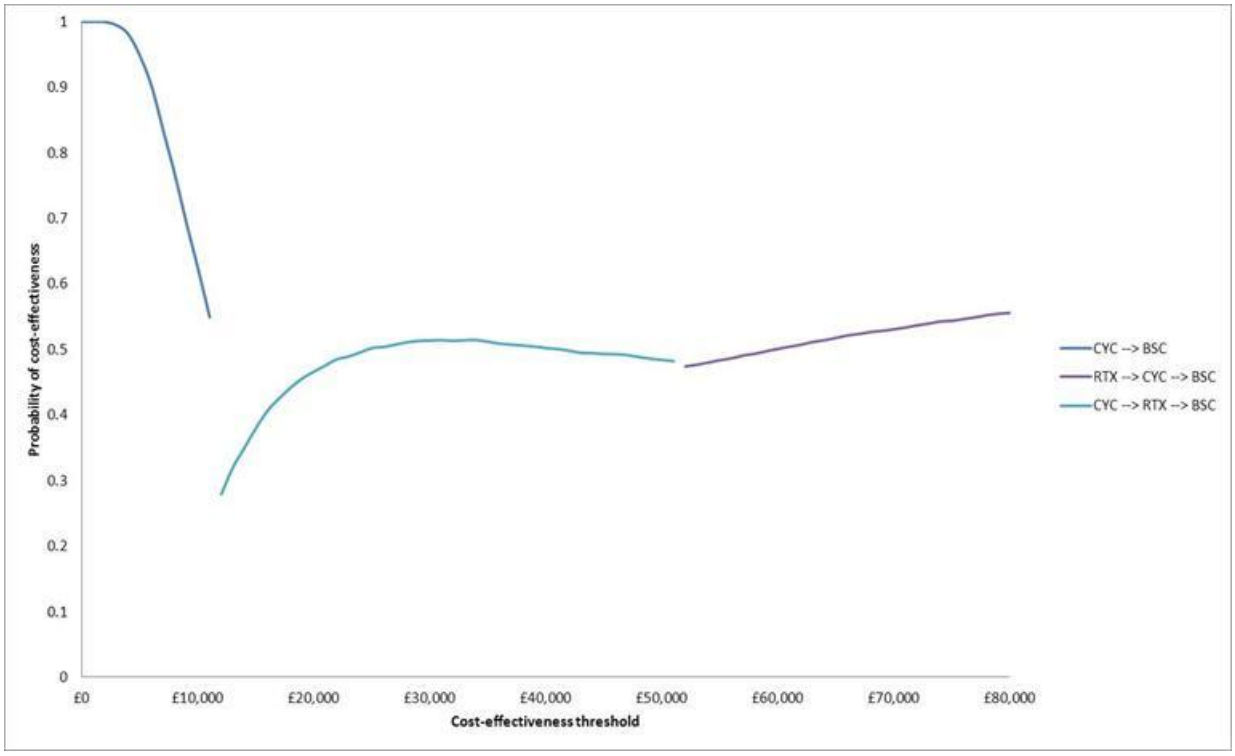


Figure 8: Cost effectiveness acceptability frontier – Recurrent disease subgroup (eligible for cyclophosphamide)



Recurrent disease subgroup – patients ineligible for additional cyclophosphamide

Table 23 presents results for the "recurrent disease" subgroup, for patients who are ineligible for further CYC treatment.

Table 23: ERG preferred analyses cost-effectiveness results – Recurrent disease subgroup (ineligible for additional cyclophosphamide treatment)

Strategy	Total Cost	Total QALYs	Inc. QALYs	Inc. Costs	ICER
BSC	£15,747.48	7.9379	-	-	-
RTX → BSC	£21,132.39	8.4412	0.5033	£5,384.90	**£10,699.45**

Table 23 demonstrates that for the "recurrent disease" analysis (in which patients are not eligible for additional CYC treatment), administering RTX treatment rather than simply best supportive care is associated with an ICER of £10,699.45 per QALY gained. This analysis is limited in that active comparators (such as MMF) are not included – instead it is assumed that best supportive care patients move directly to a low-grade, "grumbling" disease health state, in which their disease is partially controlled. CEACs and the CEAF associated with this analysis are presented in Figures 9 and 10 respectively. At a willingness-to-pay cost-effectiveness threshold of £30,000 per QALY gained, the probability that administering RTX is a cost-effective strategy in this analysis is 90.4%. The probability that not including RTX in the treatment sequence represents a cost-effective strategy is 9.6%.

Figure 9: Cost effectiveness acceptability curves – Recurrent disease subgroup (ineligible for cyclophosphamide)

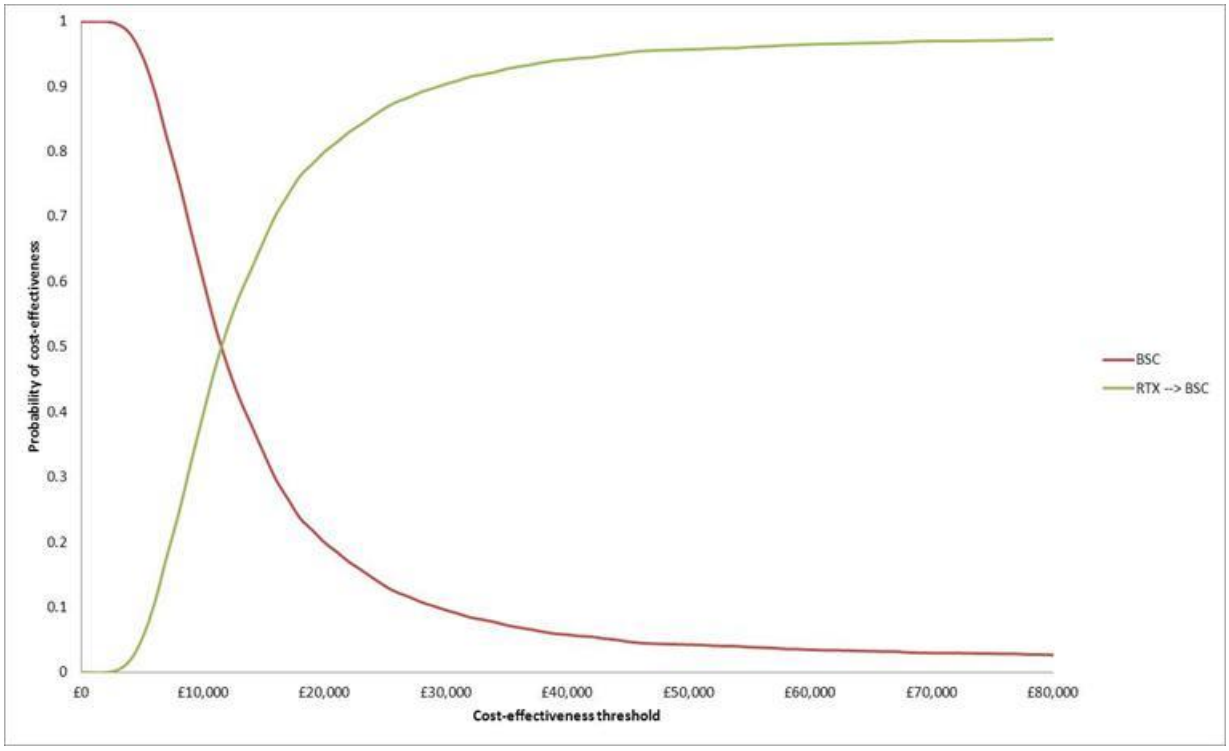
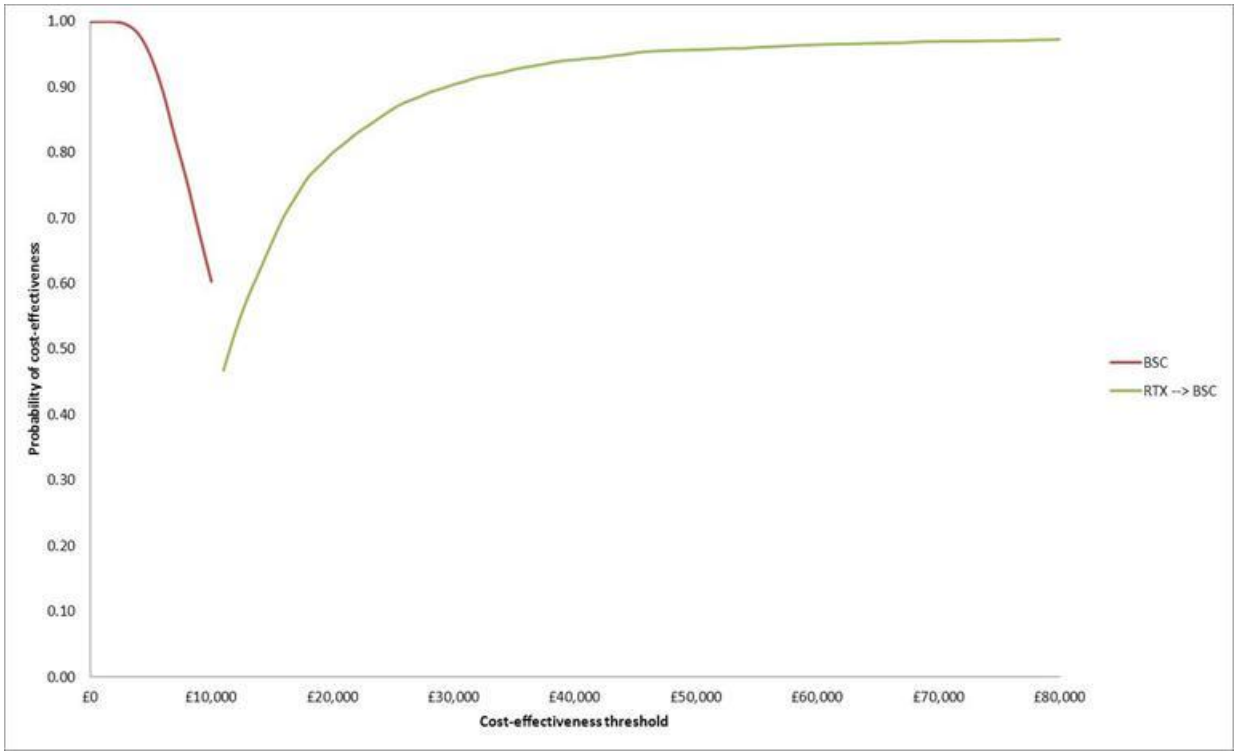


Figure 10: Cost effectiveness acceptability frontier – Recurrent disease subgroup (ineligible for cyclophosphamide)



Cyclophosphamide intolerant subgroup

Table 24 presents results for a subgroup analysis of patients who are intolerant to CYC treatment. These patients do not necessarily have recurrent disease, but cannot take CYC for some reason other than having exceeded the maximum recommended lifetime cumulative dose. Model parameter inputs are based upon the "all patient" data from RAVE.

Table 24: ERG preferred analyses cost-effectiveness results – Cyclophosphamide intolerant subgroup

Strategy	Total Cost	Total QALYs	Inc. QALYs	Inc. Costs	ICER
BSC	£15,747.48	7.9379	-	-	-
RTX → BSC	£21,184.13	8.4200	0.48	£5,436.64	**£11,277.29**

Table 24 demonstrates that for the analysis of patients who are intolerant to CYC, administering RTX treatment rather than simply best supportive care is associated with an ICER of £11,277.29 per QALY gained. This analysis is limited in that active comparators (such as MMF) are not included – instead it is assumed that best supportive care patients move directly to a low-grade, "grumbling" disease health state, in which their disease is partially controlled. CEACs and the CEAF associated with this analysis are presented in Figures 11 and 12 respectively. At a willingness-to-pay cost-effectiveness threshold of £30,000 per QALY gained, the probability that administering RTX is a cost-effective strategy in this analysis is 90.5%. The probability that not including RTX in the treatment sequence represents a cost-effective strategy is 9.5%.

Figure 11: Cost effectiveness acceptability curves – Cyclophosphamide intolerant subgroup

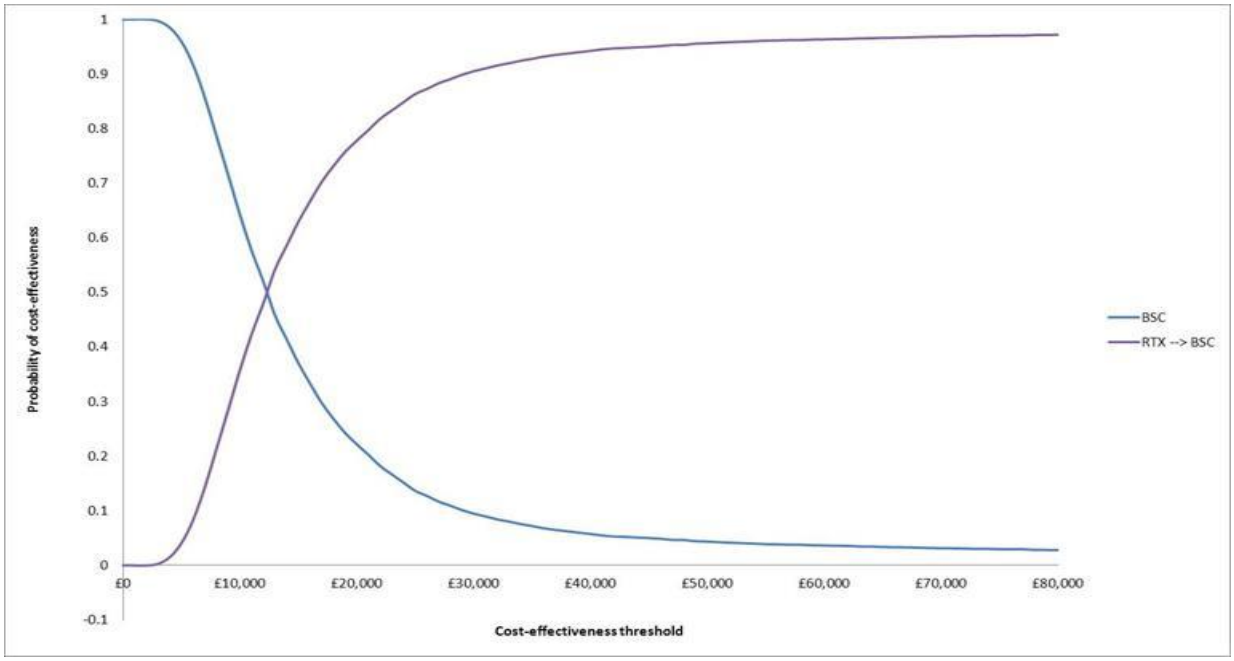
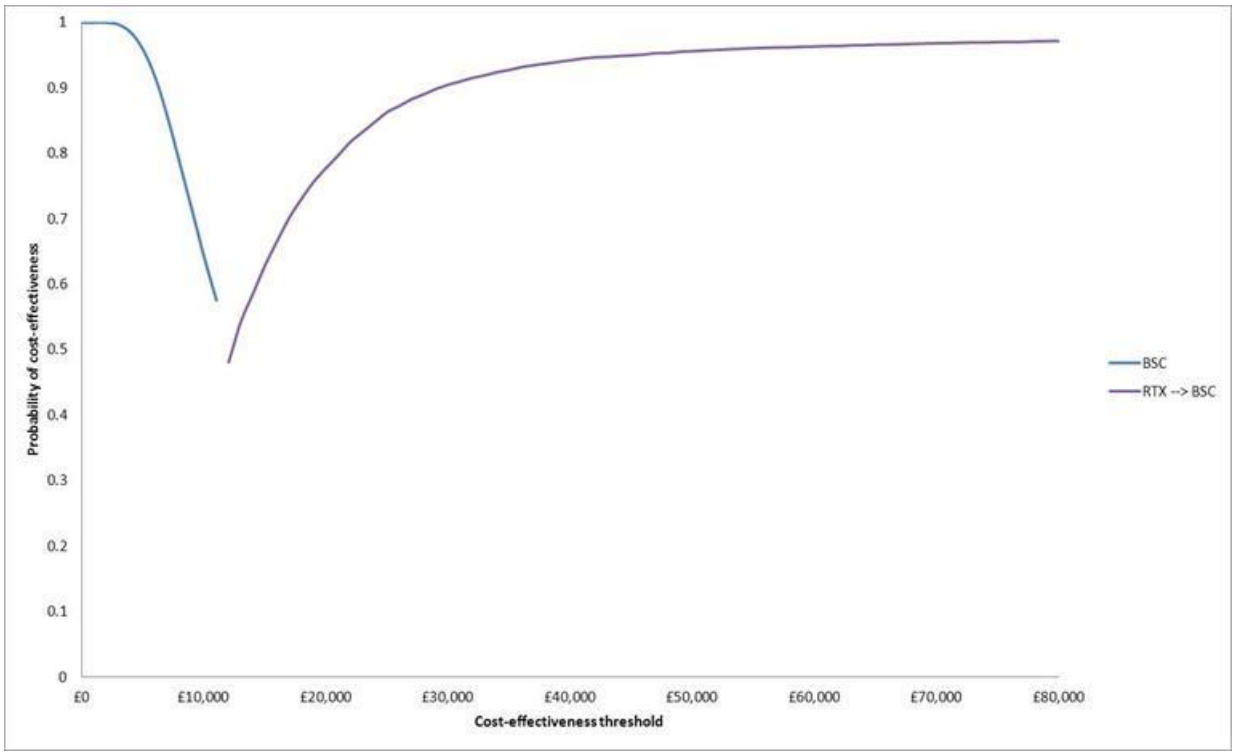


Figure 12: Cost effectiveness acceptability frontier – Cyclophosphamide intolerant analysis



6.1.4 Additional scenario analyses

Based upon comments received on the ERG's additional analyses from our clinical advisors, a set of additional scenario analyses were run. Some parameters within the amended economic model remained open to question and it was important to assess the potential impact of these on the cost-

effectiveness results. As is demonstrated in this section, these parameters had minor impacts on how the ICERs associated with different treatment sequences should be interpreted, and thus not all analyses for all subgroups were re-run. However, reporting the results of these analyses remains useful.

The parameters included the additional scenario analyses were the following:

1. **Methylprednisone administration cost**. Methylprednisone can be infused much more quickly than RTX or CYC and therefore its administration may be less costly. In our additional scenario analysis we assume its administration cost is £42.91 (rather than £180.29), based upon a consultant-led follow-up non-admitted non face-to-face outpatient appointment cost for rheumatology (currency code 410).[65]
2. **CYC administration cost**. Clinical advice received by the ERG suggests that CYC can be infused more quickly than RTX, with an infusion time of approximately half the length of the RTX infusion time. To reflect this, in our additional scenario analysis we assume an administration cost associated with IV CYC that is half the cost of RTX administration (£90.14 compared to £180.29).
3. **Trimethoprim cost**. The ERG's clinical advisors noted that in England 480mg/day of co-trimoxazole rather than trimethoprim is typically given for pneumocystis jiroveci prophylaxis. In our additional scenario analysis the cost of trimethoprim was replaced with the cost of co-trimoxazole for patients receiving CYC or RTX. This was associated with a cost of £42.81 per 6-month cycle, rather than the £21.38 associated with trimethoprim.
4. **Number of CYC administrations**. Clinical advice received by the ERG suggested that patients receiving IV CYC may often only receive 6 administrations, rather than the 10 assumed in the manufacturer's model. Typically, treatment beyond 6 administrations would only occur if evidence remained of active disease. Hence the dose and number of administrations associated with IV CYC in the economic model was amended to reflect this in our additional scenario analysis.
5. **Weight / Body Surface Area**. Although the ERG assumed a higher average weight and BSA in the analyses presented in Sections 6.1.2 and 6.1.3, our clinical advisors still felt that these may underestimate the true values for an AAV population. Data on 30 patients with vasculitis treated at Manchester Royal Infirmary provided by one of the ERG's clinical advisors gave a mean BSA of 1.90m^2 and a mean weight of 78.89kg. These are higher than the figures of 1.83m^2 and 70.51kg used in the ERG's analysis presented in Sections 6.1.2 and 6.1.3 and also, importantly, the BSA of 1.90m^2 indicates that patients treated with RTX would on average require a dose of 712.5mg per administration, thus one 500mg vial and three 100mg vials would be required, rather than the one 500mg vial and two 100mg vials assumed by the ERG in Sections 6.1.2 and 6.1.3. Based upon the distribution of BSAs observed in these 30

patients the ERG calculated that on average one 500mg vial and 2.67 100mg vials of RTX would be required per dose, and one 1000mg vial and 0.87 500mg vials of CYC would be required per dose. These figures were used in our additional scenario analysis, alongside a mean weight of 78.89kg (which affects the AZA dose).

The additional scenario analyses were conducted for the “all patients” analyses. The stepwise cumulative impact of these alterations on the ICER are presented in Table 25.

Table 25: ERG additional scenario analysis – All patients: ICERs

Strategy	ERG analysis base case	1. ↓ MP admin. cost	2. ↓ CYC admin. cost	3. Replace trimethoprim cost	4. 6 IV CYC admins.	5. ↑ BSA/ weight
CYC → CYC → BSC	-	-	-	-	-	-
CYC → CYC → RTX → BSC	£12,075.42	£11,783.47	£11,783.47	£11,829.03	£11,829.03	£12,669.64
CYC → RTX → CYC → BSC	£69,709.63	£71,265.37	£94,810.75	£94,568.00	£105,691.23	£117,545.35
RTX → CYC → CYC → BSC	£127,456.12	£128,391.53	£158,375.76	£158,229.81	£172,394.87	£191,012.75

Note: MP = methylprednisone

Table 25 shows that these amendments to the economic model have minor impacts on the ICER associated with adding RTX to the treatment sequence after CYC treatment is exhausted, with the most important effect associated with the increase in patient weight / BSA. However, the reduction in costs associated with IV CYC (through a reduced administration cost and fewer administrations), combined with the increase in RTX cost associated with a higher BSA, means that the ICER associated with moving RTX forward in the treatment sequence is associated with ICERs that are even higher than those presented in Section 6.1.3. The cumulative effect of the scenario amendments is that the ICER associated with administering RTX after one course of CYC is £117,545.35 per QALY gained compared to administering RTX after two courses of CYC, and the ICER of administering RTX as the first line of treatment is associated with an ICER of £191,012.75 per QALY gained compared to administering RTX at second line, in this "all patients" analysis. The ERG expects that the pattern in these results would be closely followed in the "treatment naïve" and "recurrent disease" subgroups.

It is noteworthy that in this analysis it is assumed that only 6 administrations of IV CYC are given. In these circumstances the cumulative dose of CYC is lower, and hence further courses of treatment may be possible. While important, the ERG does not expect that this would have an important impact upon the estimated ICERs, since one additional course of CYC would be added to each treatment sequence before CYC treatment is exhausted.

6.2 Discussion of the cost effectiveness section

The ERG undertook additional analyses to correct apparent mistakes in the manufacturer's model and also to adjust the values of model parameters that were based largely on clinical opinion. Perhaps most importantly, the ERG adjusted the manufacturer's model in order to answer two key questions:

1. Does the inclusion of RTX in the treatment sequence increase health benefits compared to the current treatment sequence?

2. If so, where is the most cost-effective place in the pathway to position RTX?

The manufacturer's model only considered placing RTX at the very beginning of the treatment pathway, although subgroup analyses were included in "recurrent disease" patients. However, the NHS Commissioning Board recommends the use of RTX as a remission induction agent in three circumstances:

Superseded – see erratum

- As an initial remission induction agent in newly diagnosed patients where avoiding CYC is desirable
- As a remission induction agent when CYC has not been effective
- As a remission induction agent at time of first relapse

Therefore, it is clearly relevant to consider the use of RTX at a later stage of the treatment sequence. This is of particular relevance given it is more costly than CYC and appears to be of similar effectiveness. For this reason, the ERG has undertaken additional work to assess where the most cost-effective place in the treatment pathway to position RTX is. Importantly, the ERG has considered scenarios in which multiple courses of RTX are *not* given, due to the lack of evidence for this. This is in contrast to the manufacturer's submission.

The additional work undertaken by the ERG indicates that including RTX in the treatment sequence increases health benefits compared to the current standard treatment sequence (that is, a treatment sequence that does not include RTX). In the "all patients" analysis, the "treatment naïve" subgroup, and the "recurrent disease" subgroup (for patients who are eligible for further CYC treatment) the ICER associated with adding RTX after CYC treatment had been exhausted was in the range of £11,129 to £12,851 per QALY gained. However, in each of these analyses the ICERs associated with administering RTX earlier in the treatment sequence were greater than £50,000 per QALY gained, sometimes substantially so. The additional scenario analysis undertaken by the ERG presented in Section 6.1.4 investigated the impact of model parameters that remained open to question. The combined impact of these amendments had minor impacts upon the ICER associated with adding RTX after CYC treatment had been exhausted – this increased to £12,670 per QALY gained in the "all patients" analysis. However, the ICERs associated with administering RTX earlier in the treatment sequence increased significantly.

In the "recurrent disease" subgroup (for patients who are ineligible for further CYC treatment) and in the "CYC intolerant" subgroup, the ICER associated with treating patients with RTX rather than best supportive care was in the range of £10,699 to £11,277 per QALY gained. In these scenarios (and in all other scenarios) "best supportive care" represents continued treatment to maintain patients in a state of low-grade "grumbling" disease. Hence these analyses are useful, but remain limited and may represent underestimates of the true ICER because relevant comparators such as MMF are excluded.

Aside from the failure to consider a full range of relevant treatment sequences in their economic model, the ERG identified several other limitations associated with the manufacturer's economic evaluation. For instance, the costs and (to a lesser extent) the utility score associated with the manufacturer's "uncontrolled disease" health state seem to be particularly prone to bias. The

manufacturer assumed that patients in the "uncontrolled disease" health state would have outpatient appointments every 1.5 weeks for the remainder of their lives – which in the manufacturer's model accounted for approximately 70% of patients' lifetimes after entry into the model. Clinical advice received by the ERG suggested that this was a vast overestimate and that in fact patients who had exhausted CYC treatment would be likely to receive other treatment in order to maintain a degree of disease control, and would be seen by clinicians much less frequently than every 1.5 weeks. In the ERG's analysis it is assumed that patients in the "uncontrolled disease" health state are treated with AZA, and have outpatient appointments once every 2 months.

This makes the ERG's analyses more clinically plausible and valid than the analysis presented by the manufacturer. The improved clinical plausibility of the ERG's analyses is strengthened by the assumption that patients who are induced into remission through RTX treatment go on to receive AZA maintenance therapy. In the manufacturer's model it is assumed that patients who achieve remission through RTX treatment do not receive any maintenance treatment – an assumption that the ERG's clinical advisors suggest is highly inappropriate. Given the ERG's assumption that AZA maintenance therapy is received, the relapse rates used within the model were also amended and were assumed to be equal in both treatment groups. The relapse rates used were based upon the severe relapse rate observed in the CYC group in the RAVE trial. This is in line with clinical advice received by the ERG that limited flares are unlikely to lead to immediate re-induction treatment – contrary to assumptions made by the manufacturer. The ERG believes that this further strengthens the clinical plausibility of the ERG's analysis in comparison to the manufacturer's analysis.

It is noteworthy that the ERG was restricted with regard to their ability to produce a fully satisfactory cost-effectiveness analysis given the limited scope of the manufacturer's economic model and the range of evidence considered therein. Clinical advice received by the ERG suggests that "partial" response to remission induction treatment represents a health state that is importantly different from "non-remission", and also that disease flares of different severity have importantly different impacts upon quality of life and subsequent treatment options. However, the manufacturer's model did not include health states for partial responders, and did not model flares of different severity. In addition, the manufacturer did not attempt to include subsequent lines of therapy with relevant treatments such as MMF, MTX, leflunomide, AZA and others in their model; for a full analysis of the relevant treatment sequences this would have been required. Hence, despite the improved clinical plausibility of the analyses presented by the ERG, important uncertainties remain. Due to being structural in nature, these uncertainties are not fully represented by the PSA presented in Section 6.1.3.

7. End of life

There are no relevant end-of-life considerations for this appraisal.

Superseded – see erratum

8. Overall conclusions

8.1 Summary of clinical effectiveness issues

The ERG consider the following limitations affecting the evidence submitted to be the principal issues affecting findings and any conclusions to be drawn from them. The MS only presents evidence on:

- An indication heavily limited by the licence (not including RTX as a maintenance treatment or for relapse other than post-CYC);
- An RTX dose that is not currently the common off-label dose in the UK;
- A single trial offering evidence on an alternative dose or regimen (RITUXVAS, RTX+CYC)
- A single trial using the dose and regimen that is to be licensed (RAVE);
- Data for only 6-12 months in the included trials, i.e. longer-term efficacy and safety outcomes are unknown;
- Some potential questions concerning certain adverse events, especially rates of mortality and malignancies.

The ERG believe that it is worthy of note that in the MS it is stated that RTX has been "*demonstrated in RCTs and clinical practice to be an effective agent in treatment of AAV … will greatly improve long term outcomes…would greatly improve the long term outcomes of patients with this… disease*" (see MS p.12).[8] The ERG believes that this overstates the case for RTX, particularly because evidence on the long-term effects of RTX have not been captured in clinical trials.

8.2 Summary of cost effectiveness issues

Based upon the ERG's version of the manufacturer's model, the incremental cost-effectiveness of adding RTX to the treatment sequence received by all patients is expected to be around £10,699 - £12,851 per QALY gained. However, for patients eligible for cyclophosphamide treatment the incremental cost-effectiveness ratio of adding RTX to the treatment sequence before cyclophosphamide is expected to be much higher – in the range of £50,842 to £317,038 for different patient groups and treatment sequences.

The ERG believes the following to represent the most important issues and uncertainties surrounding the manufacturer's submitted economic analysis:

- Several realistic treatment sequences were not modelled for the "all patients" analysis and the subgroup analyses.

- Inappropriate costs and (to a lesser extent) utilities were assumed for the "uncontrolled disease" health state (which could be more accurately described as "grumbling disease").
- An inappropriate assumption was made that all disease flares lead to immediate re-induction therapy – leading to an over-estimate of the relapse rate and unrealistically quick transition to the "uncontrolled disease" health state.
- Assumptions around the resource use costs associated with the "remission" and "non-remission" health states are questionable – the resource use assumed in the "non-remission" state in particular seems to be considerably over-estimated.
- Inappropriate assumptions were made around weight, BSA and wastage. Weight and BSA seem to be underestimated, and wastage is not included in the base case analyses.
- The manufacturer assumed that the glucocorticoid prednisone would be given alongside CYC or RTX, rather than prednisolone. In a UK context, this is inappropriate.
- The manufacturer considerably over-estimated the amount of oral CYC used in a typical treatment course.
- Several important parameters were not included in the PSA conducted by the manufacturer.

The most important of these issues relates to the treatment sequences modelled by the manufacturer. In particular the manufacturer assumed that: i) a second course of RTX would be given to patients who initially did not respond to RTX therapy; ii) patients achieving remission after RTX treatment would receive no maintenance therapy; iii) treatment sequences were not modelled in which RTX was given as an induction therapy after CYC failure. Clinical advice received by the ERG suggests that each of these assumptions is inappropriate, meaning that the decision problem modelled by the manufacturer may be misleading. The ERG amended the manufacturer's model in order to produce clinically plausible analyses. These analyses appear to demonstrate consistently that adding RTX to the standard treatment sequence is associated with an ICER of approximately £10,699 to £12,851 per QALY gained – provided RTX is only used after CYC treatment has been exhausted. This ICER increases only marginally when scenarios are run assuming a higher average BSA and weight, and assuming that patients treated with IV CYC receive only 6 cycles. Moving RTX forwards in the treatment sequence (as first- or second-line treatment in "treatment naïve" patients who are able to receive two courses of CYC, or as first-line treatment in "recurrent disease" patients who are able to receive one course of CYC) is associated with much higher ICERs. In patients who are intolerant to CYC, reluctant to take CYC for caution concerning fertility risk, or unable to take further CYC due to a high lifetime cumulative dose, the ICER associated with RTX compared to "best supportive care" is expected to be approximately £10,699 to £11,277 per QALY gained, although these analyses are limited due to the exclusion of potentially relevant comparator treatments such as MMF, and may represent underestimates of the true ICER.

Finally, it should be noted that the ERG is of the view that a great deal of uncertainty remains around the results of the ERG-amended cost-effectiveness analysis due to limitations associated with the manufacturer's model – including the failure to model varying response states, varying severities of disease flares, and the exclusion of treatments that are likely to play a part in the overall treatment pathway received by patients with ANCA-associated vasculitis. These uncertainties are not fully represented by the PSA results presented in Section 6.1.3.

8.3 Implications for research

Several trials are ongoing regarding the use of RTX for the treatment of vasculitis. These include the RAVELOS study, which is collecting long-term safety data from the RAVE trial, RITAZAREM, which is comparing RTX to AZA for maintenance of remission, and MAINRITSAN, an open label study comparing RTX to AZA for maintenance therapy at an alternative dose (see MS p.16).[8] In addition to these, the following are potentially useful areas for future research:

- The ERG assume that the RAVELOS study will provide more data on longer-term malignancies and fertility outcomes, but in addition to this longer duration studies with larger patient numbers are needed to assess these outcomes.
- Trials delivering comparative effectiveness and safety data for the 2x1g and 4x375mg/m^2 RTX dosing regimens, as well as on an RTX+CYC combination.
- Trials limited to patients with severe life-threatening disease, including:
    - Trials that include CYC intolerant patients and patients with severe life-threatening disease.
    - Trials that include RTX plus doses of CYC (as in RITUXVAS) limited to patients with severe life-threatening disease.
- Studies into long-term treatment, resource use and outcomes in patients who have exhausted RTX and CYC treatment options.
- Trials comparing RTX to other treatments such as MMF:
    - In patients who are CYC intolerant, and, potentially, in patients who are eligible for CYC treatment.
    - In patients with disease not considered to directly threaten life and without deep organ damage.

9. APPENDICES

Appendix 1: ERG searches

Table A1: Summary table of ERG searches

Database searched by ERG	Clinical effectiveness			Adverse events	
	Direct	**Indirect comparison**	**Trial registers**	**Rituximab**	**Cyclophosphamide**
Medline & Medline in Process	240	1919 (186 with RCT filter)	-	536	2696
Embase	683	3287 (463 with RCT filter)	-	1730	6479
Cochrane Library	100 (CCRCT)	66 (DARE 1 record; CCRCT 65 records)	-	61	358
ClinicalTrials.gov	-	-	512	-	-
metaRegister of controlled trials	-	-	55	-	-
WHO ICRTP	-	-	46	-	-
Total	**1023**	**5272 (715)**	**613**	**2328**	**9533**

Search strategies (Medline only)

Adapted MS search for direct evidence for rituximab and cyclophosphamide

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present

19th April 2013

1. rituximab.mp.
2. mabthera.mp.
3. rtx.mp.
4. rituxan.mp.
5. rituxin.mp.
6. cyclophosphamide.mp.
7. cytoxan.mp.
8. cyc.mp.
9. 50-18-0.rn.
10. or/1-9
11. Vasculitis/
12. vasculiti$.mp.
13. 11 or 12
14. Antibodies, Antineutrophil Cytoplasmic/
15. vasculiti$.mp.
16. 14 or 15

17. Antibodies, Antineutrophil Cytoplasmic/
18. anti neutrophil cytoplasmic antibody associated.mp.
19. anca.mp.
20. or/17-19
21. 16 and 20
22. exp Anti-Neutrophil Cytoplasmic Antibody-Associated Vasculitis/
23. 21 or 22
24. 13 and 23

Supplementary search – indirect comparisons

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present

8th April 2013

1. cyclophosphamide.mp.
2. cytoxan.mp.
3. cyc.mp.
4. 50-18-0.rn.
5. azathioprine.mp.
6. aza.mp.
7. (azasan or imuran).mp.
8. 446-86-6.rn.
9. methotrexate.mp.
10. mtx.mp.
11. mycophenol$.mp.
12. cellcept.mp.
13. or/1-12
14. Vasculitis/
15. vasculiti$.mp.
16. 14 or 15
17. Antibodies, Antineutrophil Cytoplasmic/
18. anti neutrophil cytoplasmic antibody associated.mp.
19. anca.mp.
20. or/17-19
21. 16 and 20
22. exp Anti-Neutrophil Cytoplasmic Antibody-Associated Vasculitis/
23. 21 or 22

24. 13 and 23

Supplementary search – trial registers

ClinicalTrials.gov (http://clinicaltrials.gov/)

5th April 2013

169 records for rituximab | "Autoimmune Diseases"

169 records for mabthera | "Autoimmune Diseases"

5 records for rtx | "Autoimmune Diseases"

169 records for rituxan | "Autoimmune Diseases"

0 record for rituxin | "Autoimmune Diseases"

metaRegister of Controlled Clinical Trials **(**http://www.controlled-trials.com/mrct/**)**

5th April 2013

50 records for "rituximab"

2 records for "mabthera"

2 records for "rtx"

1 record for "rituxan"

0 records for "rituxin"

WHO International ClinicalTrials Registry Platform Search Portal

5th April 2013

16 records for 14 trials found for: rituximab AND Vasculitis

14 records for 12 trials found for: mabthera and vasculitis

No results were found for: rtx and vasculitis

16 records for 14 trials found for: rituxan and vasculitis

No results were found for: rituxin and vasculitis

Conference abstracts search

Sixthteenth International Vasculitis & ANCA Workshop (2013) http://www.anca2013.com/

Link to abstracts: http://www.sciencedirect.com/science/journal/07554982/42/4/part/P2

Supplementary search – adverse events searches for rituximab

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present

5th April 2013

1. rituximab.mp.
2. mabthera.mp.
3. rtx.mp.
4. rituxan.mp.
5. rituxin.mp.
6. or/1-5
7. ((side or adverse or undesirable) adj2 (event$ or effect$ or reaction$ or outcome$)).ti.
8. (safe or safety).ti.
9. (harm$ or complication$).ti.
10. risk$.ti.
11. (treatment adj emergen$).ti.
12. tolerability.ti.
13. mortality.ti.
14. or/7-13
15. 6 and 14
16. Death/
17. Leukopenia/ci [Chemically Induced]
18. Thrombocytopenia/ci [Chemically Induced]
19. Neoplasms/ci [Chemically Induced]
20. Infection/ci [Chemically Induced]
21. Hemorrhage/ci [Chemically Induced]
22. Venous Thrombosis/ci [Chemically Induced]
23. Stroke/ci [Chemically Induced]
24. hospitalisation.ab,ti.
25. or/16-24
26. 6 and 25
27. 15 or 26

Supplementary search – adverse events searches for cyclophosphamide

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present

8th April 2013

1. cyclophosphamide.mp.
2. cytoxan.mp.

3. cyc.mp.

4. 50-18-0.rn.

5. or/1-4

6. ((side or adverse or undesirable) adj2 (event$ or effect$ or reaction$ or outcome$)).ti.

7. (safe or safety).ti.

8. (harm$ or complication$).ti.

9. risk$.ti.

10. (treatment adj emergen$).ti.

11. tolerability.ti.

12. mortality.ti.

13. or/6-12

14. 5 and 13

15. Death/

16. Leukopenia/ci [Chemically Induced]

17. Thrombocytopenia/ci [Chemically Induced]

18. Neoplasms/ci [Chemically Induced]

19. Infection/ci [Chemically Induced]

20. Hemorrhage/ci [Chemically Induced]

21. Venous Thrombosis/ci [Chemically Induced]

22. Stroke/ci [Chemically Induced]

23. hospitalisation.ab,ti.

24. or/15-23

25. 5 and 24

26. 14 or 25

Appendix 2: Critical appraisals completed by the ERG

Table A2: RAVE: Assessment using The Cochrane Collaboration's tool for assessing risk of bias

Domain	Support for judgement	Review authors' judgement
Selection bias.		
Random sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Risk of selection bias due to inadequate generation of a randomised sequence is low: "The randomization schedule will be generated, written, and controlled by Pharmaceutical Product Development, Inc. (PPDI) and will be designed to yield an assignment ratio of 1:1 between the two treatment groups within each stratum"[1]
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Risk of selection bias due to inadequate concealment of allocations prior to assignment is low: "The randomization schedule will be generated, written, and controlled by Pharmaceutical Product Development, Inc. (PPDI) and will be designed to yield an assignment ratio of 1:1 between the two treatment groups within each stratum" [1]
Performance bias.		
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Risk of performance bias due to knowledge of the allocated interventions by participants and personnel during the study is low. See efforts to minimise bias from unblinding in protocol[1], sections 3.6, 3.6.1. and 6.1
Detection bias.		
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Risk of detection bias due to knowledge of the allocated interventions by outcome assessors is low. See efforts to minimise bias from unblinding in protocol[1], sections 3.6, 3.6.1. and 6.1
Attrition bias.		
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Risk of attrition bias due to amount, nature or handling of incomplete outcome data is low: "Primary analyses were performed by the intention-to-treat" method; all participants appear to be in the analyses[2]
Reporting bias.		
Selective reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Risk of reporting bias due to selective outcome reporting is moderate to low. A 2011 update to the 2007 original protocol[3] reports a change of the primary efficacy endpoint. The change is not explained. The results for a standard

		definition of complete remission (without reference to PD dose, i.e. the secondary outcome, remission and <10mg PD at 6 months[2]) appears to demonstrate a relatively less positive outcome for the intervention than the later primary endpoint of remission with PD taper[2]. However, both sets of results are reported in the article[2], reducing the possible risk of selective reporting bias.
Other bias.		
Other sources of bias.	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	The principal concerns relate to the protocols. Two versions of the protocol exist. Clinicaltrials.gov (2007, updated 2011)[3], and the copy provided with the published article (2009)[1]. The 2009 protocol is the only source of information on randomisation and blinding, so it has had to be assumed that the planned methods were applied. The original 2007 protocol[3] records in 2011 the change of the primary efficacy endpoint. The change is not explained. The published article[2] states that subgroup analyses were "predefined" and "prespecified" but these could not be found in either version of the protocol. The trial conducts both non-inferiority and superiority analyses on the primary endpoint. The source of the non-inferiority criteria are unclear and are not justified or explained. The choice of the <10mg PD dose threshold for the secondary endpoint on complete remission is also unexplained and unclear.

PD: prednisolone

[1] *2009 NEJM protocol http://www.nejm.org/doi/suppl/10.1056/NEJMoa0909905/suppl_file/nejmoa0909905_protocol.pdf*

[2] *Stone J et al. Rituximab versus cyclophosphamide for ANCA-associated vasculitis. N Engl J Med 2010;363:221-32.*

[3] *2007 (updated 2011) clinicaltrials.gov protocol http://clinicaltrials.gov/show/NCT00104299*

Table A3: RITUXVAS: Assessment using the Cochrane Collaboration's tool for assessing risk of bias

Domain	Support for judgement	Review authors' judgement
Selection bias.		
Random sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Risk of selection bias due to inadequate generation of a randomised sequence is low: "computer minimization algorithm .. stratified by age, diagnosis and … renal function ... a 3:1 ratio for random assignment was used in view of our extensive previous experience ..." [1]
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Risk of selection bias due to inadequate concealment of allocations prior to assignment is low: ""computer minimization algorithm .. stratified by age, diagnosis and … renal function" [1]
Performance bias.		
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Risk of performance bias due to knowledge of the allocated interventions by participants and personnel during the study is high as this was an open-label study.
Detection bias.		
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Risk of detection bias due to knowledge of the allocated interventions by outcome assessors is moderate to high. "Outcomes were adjudicated by three investigators and by an independent assessor who was unaware if the study-group assignments" [1]. However, it is not clear how this outcome adjudication was achieved by these four assessors.
Attrition bias.		
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Risk of attrition bias due to amount, nature or handling of incomplete outcome data is low: ""Analyses were performed on an intention-to-treat basis" imputation by the use of the "last-value-carried forward method"; all participants appear to be in the analyses[1]
Reporting bias.		
Selective reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Risk of reporting bias due to selective outcome reporting is low. All outcomes pre-specified in protocols [2 and 3] are reported.
Other bias.		
Other sources of bias.	State any important concerns about	None

	bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	

[1] *Jones R et al. Rituximab versus Cyclophosphamide in ANCA-Associated Renal Vasculitis, N Engl J Med 2010; 363:211-220.*

[2] *ISRCTN28528813 protocol http://www.controlled-trials.com/ISRCTN28528813#?close=1*

[3] *NEJM protocol http://www.nejm.org/doi/suppl/10.1056/NEJMoa0909169/suppl_file/nejmoa0909169_protocol.pdf*

Table A4: RAVE: CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Page No and comments
Title and abstract			
	1a	Identification as a randomised trial in the title; Identification as a noninferiority randomized trial in the title	No, No
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts); See Table 2	Yes, p.221
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale; Rationale for using a noninferiority design	No
	2b	Specific objectives or hypotheses; Hypotheses concerning noninferiority, specifying the noninferiority margin with the rationale for its choice	Yes, p.222
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Yes, p.222
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	Yes, p.222-23
Participants	4a	Eligibility criteria for participants; Whether participants in the noninferiority trial are similar to those in any trial(s) that established efficacy of the reference treatment	Eligibility criteria given, p.222 and Suppl; Similarity of participants to those in trials to establish efficacy of reference treatment is not described; relevant trial is only cited in MS, and not in any other publication (WGET 2005), and population is different between these two trials
	4b	Settings and locations where the data were collected	See, Supplement only
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered; Whether the reference treatment in the noninferiority trial is identical (or very similar) to that in any trial(s) that established efficacy	Yes, p.223: Details of interventions given, but similarity to interventions in trials to establish efficacy of reference treatment is not described; relevant reference trial is only cited in MS (WGET 2005), and not in any other publication. However, WGET and RAVE CYC dose appears the same.
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed; Specify the noninferiority outcome(s) and whether hypotheses for main and secondary outcome(s) are	All outcomes clearly defined, but similarity of primary outcome to outcomes in trials to establish

		noninferiority or superiority. Whether the outcomes in the noninferiority trial are identical (or very similar) to those in any trial(s) that established efficacy of the reference treatment	efficacy of reference treatment is not described; relevant reference trial is only cited in MS (WGET 2005), and not in any other publication. The outcome of remission in the WGET reference trial is different from this trial.
	6b	Any changes to trial outcomes after the trial commenced, with reasons	Protocol reports a change in the definition of the primary outcome
Sample size	7a	How sample size was determined; Whether the sample size was calculated using a noninferiority criterion and, if so, what the noninferiority margin was	Yes, see, p.223
	7b	When applicable, explanation of any interim analyses and stopping guidelines; To which outcome(s) they apply and whether related to a noninferiority hypothesis	Yes, see Suppl.
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	See Suppl.
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	See Suppl.
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	See Suppl.
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	See Suppl.
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	Patients and outcome assessors, see Suppl.
	11b	If relevant, description of the similarity of interventions	pp.222-23
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes; Whether a 1- or 2-sided confidence interval approach was used	pp.223-24
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	pp.223-24
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Figure 1.
	13b	For each group, losses and exclusions after randomisation, together with reasons	Figure 1.
Recruitment	14a	Dates defining the periods of recruitment and follow-up	See Suppl.
	14b	Why the trial ended or was stopped	See Suppl.
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	See Suppl. and p.224
Outcomes and	17a	For each primary and secondary outcome, results for each group, and the estimated effect size	See Suppl. and p.224 and Figure 2

estimation		and its precision (such as 95% confidence interval); For the outcome(s) for which noninferiority was hypothesized, a figure showing confidence intervals and the noninferiority margin may be useful	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	See Suppl.
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	See Suppl. and pp.224-7. All analyses are described as pre-specified, but there is no published record of the analyses in any version of the protocol, only in a non-public document, an extract of which was made available to the ERG. The a priori nature of the analyses could not be confirmed.
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	See Suppl. and pp.227-8 and Table 2
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	See p.229
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	See p. 229
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence; Interpret results in relation to the noninferiority hypothesis. If a superiority conclusion is drawn for outcome(s) for which noninferiority was hypothesized, provide justification for switching	Yes, see p.229
Other information			
Registration	23	Registration number and name of trial registry	NCT00104299
Protocol	24	Where the full trial protocol can be accessed, if available	Clinicaltrials.gov
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	Yes, given, p.230 (includes funding from manufacturer of technology)

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org. Piaggio G et al. Reporting of Noninferiority and Equivalence Randomized Trials. Extension of the CONSORT 2010 Statement, JAMA. 2012;308(24):2594-2604.

10. REFERENCES

1. European Medicines Agency. Summary of opinion. *European Medicines Agency* 2013; Available from http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/000165/WC500140663.pdf (accessed May 2013).

2. Cochrane Collaboration. Risk of bias tool. 2013.

3. Piaggio G, Elbourne D, Pocock, Evans S, Altman D. Reporting of non-inferiority and equivalence randomized trials. Extension of the 2010 CONSORT statement. *JAMA* 2012; 308(24):2594-2604.

4. Falk, R.J., Jennette, J.C. Rituximab in ANCA-associated disease. *New England Journal of Medicine* 2010; 363(3):285-286. Available from http://content.nejm.org/cgi/reprint/363/3/285.pdf (accessed May 2013).

5. Garcia-Valladares, I., Espinoza, L.R. Is rituximab superior to cyclophosphamide for ANCA-associated vasculitis for induction of remission, and with a better safety profile? *Current Rheumatology Reports* 2010; 12(6):395-398.

6. Watts, R.A., Scott, D.G.I. ANCA vasculitis: to lump or split? *Rheumatology* 2012; 51:2115-2117.

7. Watts, R., Mooney, J., Skinner, J., Scott, D.G.I., MacGregor, A. The contrasting epidemiology of granulomatosis with polyangiitis (Wegener's) and microscopic polyangiitis. *Rheumatology* 2012; 51:926-931.

8. Roche Products Ltd. Single Technology Appraisal submission: Rituximab (Mabthera(R)) for the treatment of ANCA associated vasculitis. 2013.

9. Lapraik, C., Watts, R., Bacon, P., Carruthers, D., Chakravarty, K., Cruz, D. et al. BSR and BHPR guidelines for the management of adults with ANCA associated vasculitis. *Rheumatology* 2007; 46:1-11.

10. Watts, R., Lane, S., Bentham, G., Scott, D. Epidemiology of Systemic Vasculitis. A Ten-Year Study in the United Kingdom. *Arthritis & Rheumatism* 2000; 43(2):414-419.

11. Walton, E. Giant-Cell Granuloma Of The Respiratory Tract (Wegener's Granulomatosis). *BMJ* 1958;265-270.

12. Jayne, D., Rasmussen, N., Andrassy, K., Bacon, P., Tervaert, J.W., Dadoniene, J. et al. A randomized trial of maintenance therapy for vasculitis associated with antineutrophil cytoplasmic autoantibodies. *New England Journal of Medicine* 2003; 349(1):36-44.

13. Reza, M., Dornfeld, L., Goldberg, L., Bluestone, R., Pearson, C. Wegener's Granulomatosis. Long-term Followup of Patients Treated with Cyclophosphamide. *Arthritis & Rheumatism* 1975; 18(5):501-506.

14. Guillevin, L., Cordier, J.F., Lhote, F., Cohen, P., Jarrousse, B., Royer, I. et al. A prospective, multicenter, randomized trial comparing steroids and pulse cyclophosphamide versus steroids and oral cyclophosphamide in the treatment of generalized Wegener's granulomatosis. *Arthritis & Rheumatism* 1997; 40(12):2187-2198.

15. Hoffman, G.S., Kerr, G.S., Leavitt, R.Y., Hallahan, C.W., Lebovics, R., Travis, W. et al. Wegener Granulomatosis: An Analysis of 158 Patients. *Annals of Internal Medicine* 1992; 116(6):488-498.

16. Little, M.A., Nightingale, P., Verburgh, C.A., Hauser, T., de, G.K., Savage, C. et al. Early mortality in systemic vasculitis: Relative contribution of adverse events and active vasculitis. *Annals of the Rheumatic Diseases* 2010; 69(6):1036-1043. Available from http://ard.bmj.com/content/69/6/1036.full.pdf (accessed May 2013).

17. Mukhtyar, C., Guillevin, L., Cid, M.C., Dasgupta, B., de, G.K., Gross, W. et al. EULAR recommendations for the management of primary small and medium vessel vasculitis. *Annals of the Rheumatic Diseases* 2009; 68(3):310-317. Available from http://ard.bmj.com/cgi/reprint/68/3/310 (accessed May 2013).

18. NHS Commissioning Board. Clinical Commissioning Policy: Rituximab for ANCA Vasculitis, NHSCB/A3C1/1a. 2012.

19. National Institute for Health and Care Excellence (NICE). Rituximab in combination with corticosteroids for treating anti-neutrophil cytoplasmic antibody-associated vasculitis. *NICE, London* 2013; Available from http://www.nice.org.uk/nicemedia/live/13960/62550/62550.pdf (accessed Apr. 2013).

20. Roche Products Ltd. Rituximab in combination with corticosteroids for treating anti-neutrophil cytoplasmic antibody-associated vasculitis [ID567] - Roche response to clarification questions. 2013.

21. RAVE Data on File 2013.

22. Jones, R., Harper, L., Ballarin, J., Blockmans, D., Brogan, P., Bruchfeld, A. et al. A randomized trial of mycophenolate mofetil versus cyclophosphamide for remission induction of ANCA-associated vasculitis: "MYCYC". A65. *la Presse Medicale & Quarterly Medical Review* 2013; 42[4].

23. Stone, J.H., Merkel, P.A., Spiera, R., Seo, P., Langford, C.A., Hoffman, G.S. et al. Rituximab versus cyclophosphamide for ANCA-associated vasculitis. *New England Journal of Medicine* 2010; 363(3):221-232.

24. Jones, R.B., Tervaert, J.W., Hauser, T., Luqmani, R., Morgan, M.D., Peh, C.A. et al. Rituximab versus cyclophosphamide in ANCA-associated renal vasculitis. *New England Journal of Medicine* 2010; 363(3):211-220.

25. Guerry, M.J., Brogan, P., Bruce, I.N., D'Cruz, D.P., Harper, L., Luqmani, R. et al. Recommendations for the use of rituximab in anti-neutrophil cytoplasm antibody-associated vasculitis. *Rheumatology* 2012; 51(4):634-643.

26. de Groot, K., Rasmussen, N., Bacon, PA., Tervaert, J., Feighery, C., Gregorini, G. et al. Randomized trial of cyclophosphamide versus methotrexate for induction of remission in early systemic antineutrophil cytoplasmic antibody associated vasculitis. *Arthritis & Rheumatism* 2005; 52(8):2461-2469.

27. Faurschou, M., Westman, K., Rasmussen, N., de Groot, K., Flossmann, O., Hoglund, P. et al. Brief Report: Long-term outcome of a randomized clinical trial comparing methotrexate to cyclophosphamide for remission induction in early systemic antineutrophil cytoplasmic antibody associated vasculitis. *Arthritis & Rheumatism* 2012; 64(10):3472-3477.

28. Wegener's Granulomatosis Etanercept Trial (WGET) Research Group., Wegener's Granulomatosis Etanercept Trial (WGET) Research Group. Etanercept plus standard therapy for Wegener's granulomatosis. *New England Journal of Medicine* 2005; 352(4):351-361.

29. Mejia, C., Lozada, C.J. Rituximab for ANCA-associated vasculitis: A meta-analysis of randomized trials. *Arthritis and Rheumatism* 2012; 64:S660.

30. Stone, J.H., Merkel, P.A., Seo, P., Spiera, R., Langford, C.A., Hoffman, G.S. et al. Rituximab versus cyclophosphamide for induction of remission in ANCA-associated vasculitis: A randomized controlled trial (RAVE). *Arthritis and Rheumatism* 2009; 60:550. Available from http://www.blackwellpublishing.com/acrmeeting/abstract.asp?MeetingID=761&id=80182 (accessed May 2013).

31. Specks, U., Stone, J.H. Long-term efficacy and safety results of the RAVE trial. *Clinical and Experimental Immunology* 2011; 164:65.

32. Stone, J.H., Merkel, P.A., Seo, P., Spiera, R., Langford, C.A., Hoffman, G.S. et al. Extended follow-up of treatment with rituximab versus cyclophosphamide for remission-induction of anca-associated vasculitis: Which subsets are at greatest risk for flare? *Arthritis and Rheumatism* 2011; 63(10 SUPPL. 1) Available from http://www.blackwellpublishing.com/acrmeeting/abstractindex.asp?l=B&MeetingID=781 (accessed May 2013).

33. Geetha, D., Fervenza, F. The efficacy of rituximab vs cyclophosphamide for treatment of renal disease in ANCA-associated vasculitis: The RAVE trial. *Arthritis & Rheumatism* 2012; 64:S660.

34. Jones, R., Walsh, M., Jayne, D. Randomised trial of rituximab versus cyclophosphamide for ANCA associated renal vasculitis: RITUXVAS [abstract no: F-FC269]. *Journal of the American Society of Nephrology* 2008; 19:61A.

35. Jones, R.B.T. Two year follow-up results from a randomised trial of rituximab versus cyclophosphamide for generalized' ANCA-associated vasculitis: Rituxvas. *Arthritis and Rheumatism* 2010; 62:676.

36. Jones, R.B.W. Two-year follow-up results from a randomized trial of RTX versus CyP for ANCA- associated renal vasculitis: RITUXVAS. *Clinical and Experimental Immunology* 2011; 164:57-58.

37. Mukhtyar, C., Flossmann, O., Hellmich, B., Bacon, P., Cid, M., Cohen-Tervaert, J.W. et al. Outcomes from studies of antineutrophil cytoplasm antibody associated vasculitis: a systematic review by the European League Against Rheumatism systemic vasculitis task force. *Annals of the Rheumatic Diseases* 2008; 67(7):1004-1010.

38. Mansfield, N., Hamour, S., Habib, A.M., Tarzi, R., Levy, J., Griffith, M. et al. Prolonged disease-free remission following rituximab and low-dose cyclophosphamide therapy for renal ANCA-associated vasculitis. *Nephrology Dialysis Transplantation* 2011; 26(10):3280-3286.

39. de Groot, K., Harper, L., Jayne, D.R., Flores Suarez, L.F., Gregorini, G., Gross, W.L. et al. Pulse versus daily oral cyclophosphamide for induction of remission in antineutrophil cytoplasmic antibody-associated vasculitis: a randomized trial. *Annals of Internal Medicine* 2009; 150(10):670-680.

40. Walsh, M., Flossman, O., Berden, A., Westman, K., Stegeman, C., Manns, B. et al. Reduced cyclophosphamide regimens for anca-associated vasculitis are associated with an increased

risk of relapse but not death: An individual patient meta-analysis of three trials. *NDT Plus* 2010; 3:iii23.

41. Smith, R.M., Jones, R.B., Guerry, M.J., Laurino, S., Catapano, F., Chaudhry, A. et al. Rituximab for remission maintenance in relapsing antineutrophil cytoplasmic antibody-associated vasculitis. *Arthritis and Rheumatism* 2012.

42. Walsh, M., Tonelli, M., Jayne, D., Manns, B. Surrogate end points in clinical trials: The case of anti-neutrophil cytoplasm antibody-associated vasculitis. *Journal of Nephrology* 2007; 20(2):119-129.

43. Flossman, O., Bacon, P., De Groot, A.S., Jayne, D., Rasmussen, N., Seo, P. et al. Development of comprehensive disease assessment in systemic vasculitis. *Ann Rheum Dis* 2007; 66:283-292.

44. Watts R, Scott D. Recent developments in the classification and assessment of vasculitis. *Best Practice & Research Clinical Rheumatology* 2009; 23(3):429-443.

45. Chen, N., Li, X., Zhang, W., Shen, P., Yu, H., Chen, Y. et al. A randomized controlled trial of mycophenolate mofetil versus cyclophosphamide for induction therapy in ANCA associated systemic vasculitis with renal involvement. *Nephrology Dialysis Transplantation* 2012; 27:ii61-ii62.

46. Hu, W., Liu, C., Xie, H., Chen, H., Liu, Z., Li, L. et al. Mycophenolate mofetil versus cyclophosphamide for inducing remission of ANCA vasculitis with moderate renal involvement. *Nephrology Dialysis Transplantation* 2008; 23(4):1307-1312.

47. NCT00103792. Comparative Study of the Efficacy of Induction Therapy with Cyclophosphamide or Mycophenolate Mofetil for Non-life-threatening Relapses of PR3- or MPO-ANCA Associated Vasculitis. *Clinicaltrials Gov /Ct2 /Show /NCT00103792* 2004.

48. NCT00414128. A Randomised Clinical Trial of Mycophenolate Mofetil Versus Cyclophosphamide for Remission Induction in ANCA Associated Vasculitis. *Clinicaltrials Gov /Ct2 /Show /NCT00414128* 2006.

49. Jones, R.B., Ferraro, A.J., Chaudhry, A.N., Brogan, P., Salama, A.D., Smith, K.G.C. et al. A multicenter survey of rituximab therapy for refractory antineutrophil cytoplasmic antibody-associated vasculitis. *Arthritis and Rheumatism* 2009; 60(7):2156-2168. Available from http://www3.interscience.wiley.com/cgi-bin/fulltext/122466514/PDFSTART (accessed May 2013).

50. Hoffman, G.S., Drucker, Y., Cotch, M., Locker, G., Easley, K., Kwoh, K. WEGENER'S GRANULOMATOSIS. Patient-Reported Effects of Disease on Health, Function, and Income. *Arthritis & Rheumatism* 1998; 41(12):2257-2262.

51. Ndir, A., Pagnoux, C., Mahr, A., Guillevin, L., Vidal-Trecan, G. A Budget Impact Model of Rituximab Introduction in Wegener Granulomatosis Therapies. *Value in Health* 2011; 14:A411.

52. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. *NICE, London* 2008; Available from http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf (accessed 15 April 2013)

53. Office for National Statistics. Interim Life Tables, United Kingdom. *Office for National Statistics* 2011; Available from http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-223324 (accessed Apr. 2013).

54. Lane, S.E., Watts, R., Shepstone, L., Scott, D.G. Primary systemic vasculitis: clinical features and mortality. *Qjm* 2005; 98(2):97-111.

55. Ara, R., Brazier, J. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. *Value in Health* 2010; 13(5):509-518.

56. Sacco, J.J., Botten, J., Macbeth, F., Bagust, A., Clark, P. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. *PLoS ONE [Electronic Resource]* 2010; 5(1).

57. Borg, S., Glenngard, A., Osterborg, A., Persson, U. The cost-effectiveness of treatment with erythropoietin compared to red blood cell transfusions for patients with chemotherapy induced anaemia: A Markov model. *Acta Oncologica* 2008; 47:1009-1017.

58. Wang, S., Fuller, C., Choi, M., Thomas Jr, C. A Cost-Effectiveness Analysis of Adjuvant Chemoradiotherapy for Resected Gastric Cancer. *MyGCRonline Org* 2008.

59. Mathias, S., Prebil, L., Putterman, C., Chmiel, J., Throm, R., Comerota, A. A Health-Related Quality Of Life Measure In Patients With Deep Vein Thrombosis: A Validation Study. *Drug Information Journal* 1999; 33:1173-1187.

60. van den Boom, G., Rutten-van Molken, M., Molema, J., Tirimanna, P., van Weel, C., van Schayck, C. The Cost Effectiveness of Early Treatment with Fluticasone Propionate 250ug Twice a Day in Subjects with Obstructive Airway Disease Results of the DIMCA Program. *Am J Respir Crit Care Med* 2001; 164:2057-2066.

61. Sisk, J., Moskowitz, A., Whang, W., Lin, J., Fedson, D., McBean, A. et al. Cost-effectiveness of Vaccination Against Pneumococcal Bacteremia Among Elderly People. *JAMA* 1997; 278:1333-1339.

62. Erickson, P., Wilson, R., Shannon, I. Years of Health Life. *Statistical Notes* 1995; 7:1-14.

63. British National Formulary (BNF). 2013; available from http://www.medicinescomplete.com/mc/bnf/current/ (accessed Apr. 2013).

64. Barton, P., Jobanputra, P., Wilson, J., Bryan, S., Burls, A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technology Assessment, editor. 2004; 8 No 11, 1-104.

65. Department of Health. NHS Reference Costs: NHS trust reference costs schedules. 2011; available from https://www.gov.uk/government/publications/nhs-reference-costs-2009-2010 (accessed Apr. 2013).

66. Ara, R., Brazier, J. Deriving an Algorithm to Convert the Eight Mean SF-36 Dimension Scores into a Mean EQ-5D Preference-Based Score from Published Studies (Where Patient Level Data Are Not Available). *Value in Health* 2008; 11(7):1131-1143.

67. Fortner, B., Schwartzberg, L., Tauer, K., Houts, A., Hackett, J., Stolshek, B. Impact of chemotherapy-induced neutropenia on quality of life: a prospective pilot investigation. *Supportive Care in Cancer* 2005; 13:522-528.

68. Curtis, L. Unit Costs of Health & Social Care 2012. 2012; available from http://www.pssru.ac.uk/project-pages/unit-costs/2012/ (accessed Apr. 2013).

69. Chakravarty, K., McDonald, H., Pullar, T., Taggart, A., Chalmers, R., Oliver, S. et al. BSR/BHPR guideline for disease-modifying anti-rheumatic drug (DMARD) therapy in consultation with the British Association of Dermatologists. *Rheumatology* 2008; 47:924-925.

70. Chakravarty, K., et al. Quick reference guideline for monitoring of disease modifying anti-rheumatic drug (DMARD) therapy . *BSR* 2013.

71. Mukhtyar, C., Hellmich, B., Jayne, D., Flossmann, O., Luqmani, R. Remission in antineutrophil cytoplasmic antibody-associated systemic vasculitis. *Clinical and Experimental Rheumatology* 2006; 24(6 SUPPL. 43):S93-S99.

72. The European Vasculitis Society. MYCYC clinical trial protocol. *The European Vasculitis Society website* 2011; Available from http://www.vasculitis.org/Default.aspx?code=0102 (accessed Apr. 13 A.D.).

73. Latimer, N. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2011. Decision Support Unit. 13 A.D.

74. Latimer, N.R. Survival analysis for economic evaluations alongside clinical trials - Extrapolation with patient-level data: Inconsistencies, limitations, and a practical guide. *Medical Decision Making* 2013.

75. Collett, D. Modelling survival data in medical research. 2 ed. Chapman & Hall/CRC, Boca Raton; 2003.

76. Guillevin, L., Pagnoux, C., Karras, A., Khoutra, C., Aumaitre, O., Cohen, P. et al. Rituximab versus azathioprine for maintenance in antineutrophil cytoplasmic antibodies (ANCA)-associated vasculitis. *Arthritis and Rheumatism* 2012; 64:S706.

77. Jayne, D. Treating vasculitis with conventional immunosuppressive agents. *Cleveland Clinic Journal of Medicine* 2012; 79 Suppl 3:S46-S49.

78. Stone, J.H., Merkel, P.A., Spiera, R., Seo, P., Langford, C.A., Hoffman, G.S. et al. Supplementary Appendix to Rituximab versus cyclophosphamide for ANCA-associated vasculitis. *New England Journal of Medicine* 2010.

79. Jackson, R., Tosh, J., Davis, S., Wong, R., Stevenson, M., Stevens, J. Golimumab for the treatment of rheumatoid arthritis after failure of previous disease-modifying antirheumatic drugs: A Single Technology Appraisal. 2010. The University of Sheffield, ScHARR.

80. Rituximab DRAFT SmPC amendment for AAV. Annex I Summary Of Product Characteristics. *AAV* 2013.