ERRATUM

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 12th June 2013 (after check for factual accuracy)

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 <u>it would be reasonable to assume that only major relapses lead to re-induction treatment. There are limitations to this assumption, as a minority of mild relapses may result in re-treatment with CYC or RTX, but the bias associated with this assumption is likely to be less than that associated with an assumption that *all* relapses lead to retreatment.</u>

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% 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 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 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 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 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 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.

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. 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.

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 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

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. 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.

- 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 base case analysis.

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

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);
- 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.