Everolimus (Certican®) for preventing organ rejection in liver transplantation

A Critique of the Response Submitted by Novartis

Erratum

Matrix and Matrix Knowledge are trading names of TMKG Limited (registered in England and Wales under registration number 07722300) and its subsidiaries: Matrix Decisions Limited (registered in England and Wales under registration number 07610972); Matrix Insight Limited (registered in England and Wales under registration number 06000446); Matrix Evidence Limited (registered in England and Wales under registration number 07538753); Matrix Observations Limited (registered in England and Wales under registration number 05710927); and Matrix Knowledge Group International Inc. (registered in Maryland, USA under registration number D12395794).

Disclaimer In keeping with our values of integrity and excellence, Matrix has taken reasonable professional care in the preparation of this document. Although Matrix has made reasonable efforts, we cannot guarantee absolute accuracy or completeness of information/data submitted, nor do we accept responsibility for recommendations that may have been omitted due to particular or exceptional conditions and circumstances.

Confidentiality This document contains information which is proprietary to Matrix and may not be disclosed to third parties without prior agreement.
Except where permitted under the provisions of confidentiality above, this document may not be reproduced, retained or stored beyond the period of validity, or transmitted in whole, or in part, without Matrix's prior, written permission.

© TMKG Ltd, 2012

Any enquiries about this report should be directed to enquiries@matrixknowledge.com

Report commissioned by:

NHS R&D HTA Programme

On behalf of:

NICE

Produced by:

Matrix Knowledge
Peninsula Technology Assessment Group (PenTAG)

Although Optimity Matrix are primarily responsible for the work in this report, PenTAG retains responsibility for the standard of the report and the quality of the advice that it contains.

Authors:

Mariana Bacelar, Economist, Optimity Matrix
Mitesh Nakum, Assistant Economist, Optimity Matrix
Adeline Durand, Senior Economist, Optimity Matrix
Chris Cooper, Senior Information Specialist, PenTAG
Chris Hyde, Professor of Public Health and Clinical Epidemiology, PenTAG
Nick Latimer, Research Fellow in Health Economics, University of Sheffield
Jaime Peters, Research Fellow, PenTAG
Ruben Mujica-Mota, Senior Lecturer, PenTAG

Acknowledgements:

We particularly acknowledge the help of James O'Beirne, Royal Free Hospital who advised us.

Declaration of competing interest of the authors

None

Rider of 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

Bacelar M, Nakum M, Durand A, Cooper C, Mujica-Mota R, Hyde C, Latimer N, Peters J. Everolimus (Certican®) for preventing organ rejection in liver transplantation. Single Technology Appraisal NIHR HTA Programme, Matrix and Peninsula Technology Assessment Group, 2014.

Contribution of authors

Mariana Bacelar: Contributed to project management, the critique of the company's submission, report writing and editing.

Mitesh Nakum: Contributed to the critique of the company's submission, report writing and editing.

Adeline Durand: Contributed to project management, to answers the company's comments and commented on drafts of the report.

Chris Cooper: Commented on the searches provided by the company and contributed to report writing.

Chris Hyde: Contributed to the critique of the company's submission and commented on drafts of the report.

Nick Latimer: Contributed to the critique of the company's submission.

Jaime Peters: Contributed to the critique of the company's NMA.

Ruben Mujica-Mota: Contributed to the critique of the company's submission and commented on drafts of the report and is the guarantor of the report from PenTAG.

Correspondence to:

Optimity Matrix
1st Floor | Kemp House | 152-160 City Road | London | EC1V 2NP

Date completed:

January 2015

Contents

This document contains errata in respect of the ERG report in response to the manufacturer's factual inaccuracy check.

The table below lists the location of the change in the original ERG report and the nature of the change.

Page no.	Change
14	The manufacturer has requested the word 'primary' to be added to the quote taken from Barber et al. The sentence has been amended to 'Patients with alcoholic, HCV and cancer **primary** liver disease…'
14-15	The manufacturer has requested discussion around additional guidelines to be amended. The discussion around additional guidelines has been removed.
15	The manufacturer has requested the text around the NASH guidance to be amended for clarity. The text has been amended to 'However, for example, the NASH guidance mention tacrolimus monotherapy and sirolimus as relevant therapies **for non-alcoholic steato-hepatitis patients undergoing liver transplantation**.'
15	The manufacturer has requested the use of sirolimus to indicate its unlicensed use. The text has been amended to include the use of an unlicensed indication
15	Text amended 'mycophenolate acid' changed to 'mycophenolic acid'
17	As above
18	As above
18	The manufacturer has requested the text around 'Mycophenolate Mofetil' to be amended to include the company justification. The text has been amended for clarification
19	The manufacturer has requested the dosing level in the Porayko study to be more accurately specified. The following text has been amended to, 'TAC trough levels were maintained below 1.0 ng/mL **at month 12, whilst being reported to be higher in the first 4 weeks- with only limited data being reported**.'
21	The manufacturer has requested that the comment on proportion of patients dead needs to be amended for clarity. The text has been amended for clarity to '**By year 40 of the analysis 100% of patients are dead**.'
22	The manufacturer has requested the discrepancy between NMA estimates and economic model to be amended. The text has been amended to 'Any studies that evaluate the efficacy of EVR as a stand-alone intervention, would be missed by this search. **However the expected license for everolimus in liver transplantation was only in combination with reduced dose tacrolimus, any studies which might be useful for parametrisation of the model (for example) would be missed.**'
24	Text amended 'mycophenolic mofetil' changed to 'mycophenolic acid'
25	Text amended 'mycophenolate acid' changed to 'mycophenolic acid'
31	Typographical error corrected to [redacted]

39	The manufacturer has requested a sentence to be added for the use of leverage plots. The following sentence was added '**The manufacturer also provided leverage plots for BPAR in response to clarification**.'
51	Text amended, 'eexpected' changed to 'expected'
61	The manufacturer requested '[survival of liver graft]' to be removed for clarity. '[survival of liver graft]' has been removed from text
70	The manufacturer has requested the statement on the discrepancy between NMA estimates and economic model to be amended. The following amendment has been made 'Although unclear, it seems that, the estimates for the decrease in eGFR at 12 months used in the comparator arms of the economic model (Table 23) are based on the one reported for standard TAC in the NMA (Table 24).'
77	Text amended 'Sulivan et al, 2011' changed to 'Sullivan et al, 2011'
88	The manufacturer has requested the comment around 'the computational burden of the model' in accordance with the Drummond checklist to be amended. The sentence has been removed from the text.
91	Text amended 'SRS state' changed to 'SCR state'
94	The manufacturer requested '[survival of liver graft]' to be removed for clarity. '[survival of liver graft]' has been removed from text
95	Paragraph break removed
96	Text amended 'SR state' changed to 'SCR state'
97	As above
100	Text amended 'sated' changed to 'stated'
100	The manufacturer requested the health economic debate on half-cycle correction to be amended. The following paragraph regarding half-cycle correction has been removed 'Finally, from a methodological point of view, a 3 month-cycle is a relatively long one, thus a half-cycle correction should have been applied…extremely unlikely.'
101	Text amended, 'staring' changed to 'starting
101	The manufacturer requested the following text to be amended for clarity to 'By year 40 of the analysis 100% of patients are dead.'
106	Text amended, 'H23014' changed to 'H2304'
111	Sentence changed to 'However, two things should be **considered'**
114	The manufacturer has requested the dosing level in the Porayko study to be more accurately specified. The following text has been amended to, 'TAC trough levels were maintained below 1.0 ng/mL at month 12, **whilst being reported to be higher in the first 4 weeks- with only limited data being reported**.'
125	The manufacturer has requested the recommended dose of MMF as per the current SPC. The text has been amended to, '**Although the MMF SPC recommends a dose of 3000mg per day**, the ERG's clinical advisor also mentioned that the doses of oral MMF seem high after 1 year, and that 1000 or 2000 (and not 3000) mg would reflect UK clinical practice more accurately for the 1 year post transplantation period.'
132	The manufacturer requested the discussion around PSA to be amended for

	clarity. The paragraph has been amended

2.0 Background

2.1 Critique of company's description of underlying health problem

In Sections 2.1 to 2.3 of their submission, Novartis describe the underlying health problem. They provide a summary of the characteristics and progression of end stage liver disease (ESLD). It is mentioned how the most common underlying causes of ESLD in Europe are hepatitis C virus infection (HCV), nutritional-toxic liver cirrhosis (NTLC) and hepatocellular carcinoma (HCC). It is also explained how patients with ESLD rely on liver transplantation to survive as there are is no suitable alternative treatment for these patients.

The complications associated with liver transplantation are summarized. The company explains that in the short-term there is a risk of acute rejection of the graft which can progress to chronic rejection if the immunosuppressive therapy is not adequately managed. Long-term complications are typically associated with the recurrence of the underlying liver disease or the immunosuppressive regimen used (these typically include renal dysfunction, new onset of diabetes mellitus and cardio-vascular disease).

The company provide information on the median survival time of adult liver allograft recipients in the UK (based on Barber et al, 2007). The estimated median survival is 22.2 years (CI 19.3 - 25.6, p-value 0.05) and an estimated loss of 7 years compared with age and sex matched population. The median survival for specific age groups is also provided, with patients between 24 and 54 years reporting a median survival of 25.3 years (CI 20.5–31.2) and 55 to 64 year old patients, reporting a median survival of 19.5 years (CI 16.1–23.6).

To note is that Barber et al, 2007 also reports the median survival for transplant patients with specific underlying liver disease. Patients with alcoholic, HCV and primary cancer liver disease had a median survival of 15, 12 and 5.3 years, respectively.

Novartis present the incidence of liver transplants from April 2012 – March 2013 based on a reliable relevant source (NHS Blood and Transplant). However, the NHS Blood and Transplant present more recent figures in their website from 1 April 2013 to 31 March 2014, which the ERG reproduce here.

Table 1. Updated incidence of liver transplants in the UK

Description	UK estimate	Source
Liver transplants in the UK (including deceased and living transplants) April 2013 – March 2014	881	http://www.organdonation.nhs.uk

2.2 Critique of company's overview of current pathway of care and service provision

Novartis mention that existing NICE guidance and national protocols do not cover the management of immunosuppressive treatment for all patients who undertook liver transplant.

Furthermore, on page 135 of their submission Novartis explain that the economic analysis considered patients transplanted for any kind of liver failure, such as NASH.

Clinical opinion sought by the ERG explained that there is huge variance in clinical practice with regards to post-transplantation immunosuppressive treatment regimens in the UK. It also agreed that the only available guidance on the management of immunosuppressive treatment for patients who undertook liver transplant is the BTS NASH guidance and deemed this extremely relevant for the context of the decision problem.

The company explains that everolimus will fit in the existing post-transplantation clinical pathway as it offers an alternative to current immunosuppressive therapy. It is also mentioned that current practice shows variations with regards to choice of immunosuppressive therapy and long-term treatment to ensure graft survival whilst avoiding treatment-related complications.

On page 33, it is reported that calcinerium inhibitors (CNIs) remain the *backbone of immunosuppression* for post-transplanted liver patients regardless of the well-known long-term complications associated with these, such as renal toxicity. Therefore there seems to be a delicate balance between the use of these to reduce acute rejection and trying to reduce CNI-related complications in the long-term. Because everolimus acts synergistically with CNIs, there is an opportunity to minimise the CNI dose administered to patients after liver transplant, therefore sparing them from some of the complications associated with long-term use of CNIs.

The company point to the fact that in the UK clinical practice is based on tacrolimus-containing regimens for the prophylaxis of hepatic graft rejection.

There is not much consideration given to the main immunosuppressive therapies used in the UK apart from mentioning there is variance in practice and that most regimens contain tacrolimus in combination with other drugs. However, for example, the NASH guidance mention tacrolimus monotherapy and sirolimus as relevant therapies for non-alcoholic steato-hepatitis patients undergoing liver transplantation.

Clinical opinion sought by the ERG advised that about 10-15% of patients currently receive sirolimus, although the therapy is unlicensed, which like everolimus is an mTOR inhibitor, instead of everolimus during their first year of immunosuppressive therapy. Use of sirolimus is significantly lower than everolimus since the drug is poorly tolerated and has been known to cause hepatic artery thrombosis. Everolimus is much better tolerated, and can be introduced in the immunosuppressive therapy earlier than sirolimus.

Furthermore, the ERG's clinical advisor explained that around 30% of patients receive tacrolimus monotherapy, which will be adjusted to lower doses as the patient gets stable and that around 70% of patients receive therapies including 2 agents, like mycophenolic acid or azathioprine in combination with reduced doses of tacrolimus.

EVR was submitted to the MHRA on the 27th May 2014 and UK approval is anticipated for early December 2014. The anticipated UK indication for liver transplantation is f*or the prophylaxis of organ rejection in patients receiving a hepatic transplant. In liver transplantation, Certican should be used in combination with tacrolimus and corticosteroids.* The company also expect Certican to be assessed by the SMC, with a submission estimated for [redacted] and with guidance expected to be published in [redacted].

The company reports that the EVR recommended starting dose is 1.0 mg twice daily, 4 weeks after transplantation. The EVR dose is targeted to remain at a trough level of 3 to 8 ng/mL in combination with reduced TAC at 3 to 5 ng/mL and corticosteroids, which can be adjusted if necessary.

Whilst for EVR the reported doses are in line with the treatment regimen followed in the main clinical trial (H2304), source of effectiveness evidence submitted by the company, this is not the case for reduced TAC, where trough levels in the trial were above 5. Of note is that in the UK, a reduced dose of TAC is defined as blood through levels of TAC <5 ng/mL. This will be further discussed throughout the ERG report.

The company does not specify the corticosteroid recommended dose, however in H2304 this was given at a minimum dose of 5mg per day for 6 months.

The company reports that dose adjustments might be necessary. The standard EVR dose should be halved for patients with moderate or severe hepatic impairment (Child-Pugh B or C).

Treatment is expected to be continued throughout the remaining of the graft/patient life.

3.3 Comparators

The comparator set out in the scope was defined as a standard immunosuppressive therapy with a CNI (such as ciclosporin or tacrolimus) and a corticosteroid in combination with azathioprine or mycophenolic acid.

Novartis report that the choice of relevant comparators was based on historical discussions with NICE and a scoping workshop (held on the 24th July 2013) which resulted in the selection of tacrolimus in combination with mycophenolate mofetil with or without corticosteroids (hereafter TAC+MMF) and azathioprine in combination with tacrolimus, with or without corticosteroids (hereafter TAC+AZA) as the relevant comparators.

Thus the comparator used was TAC+AZA and TAC+MMF, both with corticosteroids for the initial 6 months of therapy.

Therefore, the comparator used in the submission departs from the one set out in the initial scope in three aspects:

1. It excludes ciclosporin from the economic analysis
2. It considers mycophenolate mofetil (MMF) instead of mycophenolic acid
3. It considers drug regimens both with and without concomitant corticosteroids

Clinical opinion sought by the ERG indicated that approximately 7% of patients have ciclosporin included in their treatment regimen but that this drug would not be considered as a first-line treatment option due to its high rate of adverse events, therefore the exclusion of ciclosporin seems to be appropriate.

The scope originally considered mycophenolic acid. However the company decided to define mycophenolate mofetil as the comparator drug. Clinical opinion sought by the ERG explained that mycophenolic acid includes both mycophenolate sodium and mycophenolate mofetil. Even though the active component of mycophenolate mofetil and mycophenolate sodium is the same, these drugs are not interchangeable as 500mg of mycophenolate mofetil is equivalent to 360mg of mycophenolate sodium.

Furthermore the justification provided by Novartis on page 36 of their submission regarding the choice of mycophenolate mofetil is not clear to the ERG, however the company stated in the clarification document that mycophenolate sodium does not have a licence for use in liver transplantation. Therefore it seems that mycophenolic acid should have been used as the comparator drug as it is less restrictive than mycophenolate mofetil or mycophenolate sodium.

As previously mentioned, the exclusion of corticosteroids after 6 months is appropriate as only patients with specific co-morbidities such as autoimmune hepatitis will be kept on corticosteroids after 6 months (as per clinical opinion sought by the ERG).

Due to a wide range of studies considered for the NMA and time constraints, the ERG approach was to focus on:

- The studies included in the final NMA (Novartis needed to refine the studies included in the analysis due to incompatibility reasons)
- The studies used to derive the parameters included in the economic model (tBPAR and renal functioning)

Nevertheless, the ERG are still not clear which studies have been included in the final NMA analysis for the biopsy proven acute rejection (tBPAR) outcome, due to lack of clarity and transparency in the submission.

The studies included in the NMA to obtain renal outcomes were the H2304 trial, Neuberger et al, 2009, Boudjema et al, 2011, Porayko et al., 1994 and McDiarmid et al., 1993. The doses for the comparator drugs are presented in Table 2.

Table 2. Drug doses in comparator studies

(Neuberger *et al.*, 2009[65])	Standard TAC	The standard TAC arm presents average trough levels of 11 ng/mL until month 3 and then decreased to 9 by the end of the 1st year.
	MMF + reduced TAC	The reduced TAC arm presents average trough levels of 9 ng/mL until month 3 and then decreased to 8 by the end of the 1st year. MMF, 2g/day

(Boudjema *et al.*, 2011[83])	Standard TAC	The standard TAC arm presents average trough levels above 10 ng/mL for the whole of the 1st year. The reduced TAC arm presents average trough levels between 10 and 9 ng/mL for the 1st year. MMF, 1.5 g twice a day for the first 6 weeks, 1.0 g twice a day until month 12.
	MMF + reduced TAC	
(Porayko *et al.*, 1994[88])	Standard TAC	TAC trough levels were maintained below 1.0 ng/mL at month 12, whilst being reported to be higher in the first 4 weeks- with only limited data being reported. AZA: 2mg/kg/day.
	AZA+ciclosporin	
(McDiarmid *et al.*, 1993[93])	Standard TAC	TAC trough levels were maintained below 2 ng/mL. AZA, 1-2mg/kg/day.
	AZA+ciclosporin	

There are three additional points which arose from discussions with our clinical expert which the ERG believe to be worthy of consideration:

- When evaluating the benefit of EVR as a concomitant drug prescribed with a reduced dose of TAC and with regards to its renal sparing effect, EVR is likely to perform as good as MMF and AZA when given with reduced doses of TAC. This is because the renal sparing effect comes from allowing reduced doses of TAC (versus standard doses of TAC), therefore using EVR, MMF or AZA is likely to lead to the same renal outcomes. In this case, the cheaper drug is likely to be the most cost-effective therapy. That would be AZA.

- The true advantage of EVR and the uniqueness of the drug is likely to be the fact that it can be used as a monotherapy regimen. The ERG are aware that the drug indication is not monotherapy, nonetheless we consider this to be an extremely relevant issue. Whilst MMF and AZA cannot be given as monotherapy regimens (as it has been shown this leads to an increase in graft loss), EVR seems to perform well when taken alone. Disappointingly, it seems that not enough data exist to show the benefit of EVR monotherapy. Even though the H2304 trail initially designed a TAC discontinuation arm, where patients were kept on EVR monotherapy, this arm was discontinued before the end of the trial, as it reported higher rates of episodes of acute rejections. However our clinical advisor explained that this is likely to be related to the inappropriate choice of clinical endpoints for the trial which focused on episodes of acute rejection instead of long-term survival of the graft. Nowadays it seem to be broadly accepted that patients can experience (up to 2) episodes of acute rejection without these impacting on the long-term survival of the graft as these are easily treated and have around 90% of successful resolution (also demonstrated in Novartis' economic analysis). Therefore, whilst the number of acute rejections is a relevant endpoint, long-term survival would be a better one for clinical trials evaluating the effectiveness of immunosuppressive therapies.

Clinical opinion sought by the ERG informed that graft function and fibrosis are difficult to measure as their definition is not standardized. Additionally, the impact of immunosuppressive therapy on graft function and fibrosis would only become apparent around year 3 of treatment. The same applies to the recurrence of HCC, which would only become apparent after 2 years of therapy. Therefore the ERG's clinical advisor explained that the main clinical trial H2304 (which lasted for 1 year with a follow-up of 36 months) would not necessarily capture the impact of EVR+rTAC on these outcomes. However, it was mentioned that if any impact was to be observed it was likely to benefit EVR+rTAC (against standard TAC).

3.5 Time frame

The time horizon considered in the economic model was lifetime (80 years). Given that the average staring age of patients in the economic model was 54 years (with a standard deviation of 10 years), the time frame seems unnecessarily high, adding to the computational burden of the model. By year 40 of the analysis 100% of patients are dead.

4.0 Clinical effectiveness

4.1 Critique of company's approach

In this chapter we assess the clinical evidence provided by Novartis in their submission.

We start with a description and critique of Novartis's literature search strategy, followed by a description of the main studies selected for clinical effectiveness and their quality assessment. We then look at the company's selection of outcomes and the statistical approach they used to analyse them. This is followed by a summary of their submitted evidence for clinical effectiveness and our comment on their validity.

4.1.1 Description of company's search strategy and comment on whether the search strategy was appropriate

4.1.1.1 Clinical Effectiveness Searches

The ERG are happy to accept the clinical effectiveness searches as presented by the company.

Novartis ran four systematic literature reviews in order to identify relevant published and unpublished clinical data. These targeted:

1. Clinical data for the intervention of interest, EVR+rTAC
2. Studies suitable for a network meta-analysis (NMA), both for the intervention and the two comparator regimens (AZA+TAC and MMF+TAC) with or without corticosteroids
3. Non-RCT data for the intervention of interest, EVR+rTAC
4. Adverse event (AE) data

The search strategy was last updated in August 2014, so the results are considered current for this submission. The effectiveness search syntax took the following form:

(Terms for liver transplant or hepatic transplantation or graft) AND (terms for Everolimus (including brand names Certican or Zortress)) OR (terms for Azathioprine) OR (terms for mycophenolic acid) AND (terms for cyclosporine) OR (terms for tacrolimus).

Novartis searched all of the required bibliographic databases, in addition to clinical trial registries and conference proceedings. The ERG are content with the range of resources used in this submission and, therefore, the company's attempts to locate published and unpublished RCT evidence.

The ERG point to the following limitations of the searches undertaken:

The search returns are limited to studies that use cyclosporine or TAC in combination with EVR or AZA or MMF. Any studies that evaluate the efficacy of EVR as a stand-alone intervention, would be missed by this search. However the expected license for everolimus in liver transplantation was only in combination with reduced dose tacrolimus, any studies which might be useful for parametrisation of the model (for example) would be missed.

The bibliographic searches were date limited 1990-Current and the conference proceedings were date limited 2012-current.

The ERG noticed a small typographical error in the clinical effectiveness searches. The Boolean connecter OR had been inadvertently omitted between mycophenolic acid / morpholinoethyl ester. The line was presented in the company's submission as:

(cellcept or mycophenolic acid morpholinoethyl ester or RS 61443 or RS-61443 or mycophenolate mofetil or Mycophenolate mofetil hydrochloride or MMF).mp.

In clarification, the company confirmed it should read:

(cellcept or mycophenolic acid or morpholinoethyl ester or RS 61443 or RS-61443 or mycophenolate mofetil or Mycophenolate mofetil hydrochloride or MMF).mp.

Novartis have provided the ERG with a list of the 7 unique studies resulting from correcting this error. These studies were double-screened and all 7 studies were excluded. The ERG are content that this point has been dealt with satisfactorily.

- ### Adverse events

The company used their clinical effectiveness search strategy to identify studies reporting adverse events. This strategy worked as the company did not limit their clinical effectiveness searches by study design (i.e. to RCTs using an RCT search filter). The ERG are happy to accept these searches.

Outcomes	• ***Time to recurrence of hepatocellular carcinoma***[2] *;* • *Renal function;* • *Time to end-stage renal disease;* • *Adverse effects of treatment;* • *HRQL.*	• *Studies that do not focus rejection of the liver as an outcome (efficacy) or HRQL.* • *Studies with only cost and no clinical outcomes.*	*These include all the outcomes derived from the final scope issued by NICE, except for the following:* • *Graft function / fibrosis* • *Expert opinion advised that in clinical experience there was limited evidence to differentiate between interventions with regard to the two outcomes above.*
Study design	• *RCTs of any duration, including cross-over RCTs if data were presented at cross-over.* • *Studies published as abstracts or conference presentations were eligible for the primary analysis of clinical effectiveness if adequate data are provided.*	• *Non-RCT study designs or articles reporting results of RCTs published elsewhere, e.g. reviews, meta-analyses/pooled analyses, editorials, notes, comments or letters.*	• *Only RCTs were considered in line with the objective of this literature search.*

Source: Submission Table 4

To note is that for the NMA the company decided to develop a refined criteria as to include any study that included two or more of the following comparators within the study:

1. *Everolimus plus reduced dose tacrolimus with or without a corticosteroid.*
2. *Any combination of MMF and a calcineurin inhibitor (reduced/standard ciclosporin or reduced/standard dose tacrolimus monotherapy) with or without a corticosteroid.*
3. *Any combination of azathioprine and a calcineurin inhibitor (reduced/standard dose ciclosporin or reduced/standard dose tacrolimus monotherapy) with or without a corticosteroid.*
4. *Tacrolimus monotherapy with or without corticosteroid.*

The scope defined the intervention as EVR in combination with TAC and a corticosteroid, however the decision problem addressed in the submission looked at the use of EVR+rTAC with or without corticosteroids. The specification of reduced TAC seems appropriate in theory as the indication of EVR is in combination with a reduced dose of TAC. However, as previously mentioned the reduced TAC dose in H2304 is the equivalent to a standard dose of TAC in UK practice.

After a clarification request from the ERG with regards to the inclusion of ciclosporin in the undertaken literature reviews, the company stated that *all relevant comparators AZA or MMF in combination with a calcineurin inhibitor (ciclosporin or TAC) with or without corticosteroids* were included in all the literature searches. However, after consultation with clinical experts, the company decided to exclude ciclosporin from the economic analysis. This is appropriate as previously explained by the ERG.

After clarification, the company confirmed that mycophenolate mofetil was used in the submission instead of mycophenolic acid, originally defined in the scope. The company claims that there are two

presentations of mycophenolic acid (mycophenolate mofetil and mycophenolate sodium) but that mycophenolate sodium is not licensed for use in liver transplantation (only for renal transplantation). Clinical opinion sought by the ERG informed that in the UK most clinicians will prescribe mycophenolic acid and that even though mycophenolate sodium and mofetil cannot be interchangeably used, the active component is the same and only a dose adjustment is necessary as 500mg mycophenolate mofetil = 360mg of mycophenolate sodium. Therefore, the ERG do not see a valid reason to specify the type of mycophenolate in the inclusion criteria, and the broad mycophenolic acid term could have been used instead.

Novartis clarified that time to recurrence of hepatocellular carcinoma was included in the searches while graft function and graft fibrosis were omitted as per expert opinion.

The submission includes a flow diagram that shows the number of studies identified through the database searches and the number of studies included and excluded at each stage of the review and the reasons for exclusion.

Overall, the inclusion criteria seems appropriate to identify all the relevant evidence set out in the NICE scope.

4.1.3 Studies included and excluded

The search strategy identified one RCT, H2304 which studied the intervention of interest (EVR+rTAC with or without corticosteroids) and 33 individual records which were related to this RCT. These can be found in Table 5 (page 44) in the submission.

To note is that the company use different references to quote H2304 throughout the submission as some of them refer to papers, posters or presentations. H2304 is a phase III RCT for which there are 3 main clinical study reports (CSRs), corresponding to a 12-month, 24-month and a 36 months analysis to evaluate the long-term efficacy and safety of concentration-controlled EVR in liver transplant recipients. The main references for these are:

- Hexham et al, 2011 – CSR for the 12-month analysis
- Lopez et al, 2013 - CSR for the 24-month analysis
- Rauer et al, 2014 - CSR for the 36-month analysis

A brief description of the H2304 trial is given in Table 4 below. The company mainly quotes de Simone et al, 2014 as the reference for H2304. This is a study funded by Novartis, looking at the 12-month data from Hexham et al, 2011 CSR.

Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	*All randomised patients were included in the intent-to-treat population. This was appropriate but the methods used to account for missing data are unclear.*	*Yes*	Again, the ERG find it strange that the company are not clear about the methods used to account for missing data given that H2304 is a Novartis trial.

Response source: Novartis Submission Table 98

Novartis have covered the elements used in the critical appraisal of RCTs according to the Centre for Reviews and Dissemination Systematic Reviews checklist (2008).

Additionally, the ERG note that a limited number of UK patients were involved in the trial. The study was conducted predominantly with patients in the US, with only [redacted] in the UK.

Novartis assessed the NMA studies for their validity (Table 16 in the submission). Even though they follow the template suggested by NICE to assess the NMA studies, the ERG do not find that it provides very useful information as all the questions were answered with a Yes/No/Unclear reply. So for example, one of the criteria used by Novartis to assess the validity of the trials is "Were the groups similar at the outset of the study in terms of prognostic factors?". If the company considered this to be a "Yes", there is no way for the ERG to assess this answer as the submission does not state the criteria used to assess similarity in terms of prognostic factors.

Also the proportion of answers answered with "unclear" was considerably large. For example, for the first NMA question, 'Was randomisation carried out appropriately?' 68% of the answers were answered "unclear". Overall for all questions answered, 40% were answered "Unclear".

Studies were also grouped into categories within the critical appraisal section, so it was difficult to assess how each study had been individually appraised.

4.1.6 Description and critique of company's outcome selection

- H2304 study

Primary Efficacy Endpoint

The primary efficacy end-point in H2304 was a composite of 3 outcomes defined in the scope. These are the efficacy failure rate of treated biopsy proven acute rejection (tBPAR), graft loss or graft death at 12 months post transplantation (excluding any events before randomisation).

tBPAR was determined as acute rejection with a locally confirmed rejection activity index (RAI) ≥ 3 according to the Banff 1997 criteria when treated with anti-rejection therapy. The Banff RAI includes 3 components scored from 0 to 3: venous endothelial inflammation (E), bile duct damage (B), and portal inflammation (P). The scores are combined to an overall score (the RAI).

submission or from a theoretical exercise, as the data included in the codes do not relate to the submission data in any way. This poses a major limitation as the ERG could not verify which data were used for the analysis of specific outcomes. This represents a major concern as the ERG are not clear which studies have been included in the NMA analysis for the tBPAR outcome, due to lack of clarity and transparency in the submission.

The NMA was conducted using a Bayesian framework. The WinBUGS/OpenBUGS software package was used to estimate the parameters of the different NMA models using a Monte Carlo Markov Chain (MCMC). Likelihood and link functions were defined for the different types of outcome data. Non-informative prior distributions were used for model parameters,

A non-informative prior distribution was used for the model parameters. Nuisance and treatment effect parameters followed a normal distribution, μ_{jb}~normal (0, 10,000) and d_{Ak}~normal (0, 10,000) respectively. The heterogeneity parameter was based on a uniform distribution, σ ~ Uniform (0, 5).

The values attributed to the normal distributions were in line with what is suggested in NICE DSU guidance (Dias et al, 2011), however the between-trial variance parameter, σ ~ Uniform (0,5) presents an extremely high upper limit. The ERG requested clarification as to why the upper limit was attributed the value of 5, however Novartis have not clarified this. The fact that this parameter has a high value might be leading to some of the considerably large credible intervals reported for the NMA results.

Both random and fixed effects models were ran with the latter being preferred for most outcome measures. The company presented the Deviance Information Criterion (DIC) and upon the ERG request, the posterior mean of the deviance was also presented to ensure that the selected model's overall fit was adequate. These are presented in Table 8.The manufacturer also provided leverage plots for BPAR in response to clarification

In NMAs, convergence of results is normally assessed by using the Brooks-Gelman-Rubin diagnostic in WinBUGS and the Monte Carlo error, which reflects both the number of simulations and the degree of autocorrelation. However Novartis only reported an initial series of iterations which were discarded as 'burn-in' and inferences based on additional iterations. Therefore there was not enough information available to reach conclusions regarding convergence of results'

Table 8: Model fit statistics

Outcome Measure	Fixed Effects			Random Effects			Model selected
	DIC	Dbar	pD	DIC	Dbar	pD	
Overall Survival 12 months	133.3	115.3	17.9	134.8	115.3	19.5	FE
Overall Survival 24 months	83.2	71.4	11.8	85.0	71.8	58.6	FE
Graft survival 12 months	127.8	97.8	15.9	129.6	98.5	16.9	FE
Graft survival 24 months	86.3	74.3	11.9	88.2	75.1	13.1	FE
tBPAR 3 months	138.3	121.4	16.9	139.9	120.9	18.9	FE*
tBPAR 6 months	79.9	68.9	11.0	79.5	67.4	12.0	FE*
tBPAR 12 months	119.7	105.6	14.1	119.3	102.5	16.9	FE*
Renal function (eGFR) 12 months	-0.8	-3.8	3.0	0.3	-4.2	4.4	FE

The data ultimately used in the economic model, didn't consist of OR but instead on probability of events. The ERG requested that the company explained how these were derived, to which the company replied by presenting what was considered *the absolute results for the NMA outcomes.* These were basically the final values used in the model, for which the ERG requested initial clarification. The ERG are aware that these absolute estimates can be obtained through a WinBUGS command. Nonetheless, in order to be able to compute these in the software package, baseline effectiveness data need to be inputted for the comparator treatments. Novartis do not provide these data in their submission, nor make reference to it. Again, because the ERG didn't get access to the WinBUGS code using the actual submission data, we could not verify this.

The absolute NMA results are now presented.

Absolute estimate results for NMA

The ERG present results for the following estimates:

- Overall survival at 12 and 24 months
- Graft survival at 12 and 24 months
- tBPAR free at 3, 6 and 12 months
- Expected absolute results for renal function at 12 months

However, as only the probability of being tBPAR free and renal function were used in the economic model, we mainly focus on these outcomes.

The company also presented the expected absolute results for withdrawals due to adverse events at 12 months (hypertension and diabetes) and the expected absolute results for infections and HCV recurrent 12 months. The reader should consult the clarification document for details on these.

For overall survival and graft loss, Novartis have used different studies to report the 12 month analysis and the 24 month analysis. This was also the case for time BPAR free, where studies used in the network varied for 3,6 and 12 months. Due to incompatibilities, only studies reporting a 12 and 24 month analysis for overall survival and graft loss were used. The same was applied to time BPAR free, where only the studies reporting a 3, 6 and 12 month analysis were used with goal to increase compatibility.

The ERG interpreted that by "compatibility" the company meant the decreasing trend observed in results across the various time frames. Novartis reported that the expected absolute point estimates for overall survival at 12 and at 24 months were compatible for EVR + reduced TAC (87.1% [95%CrI: 71.1; 94.9] at 12 months and 85.3% [95%CrI: 72.5; 92.9] at 24 months), but that incompatibility issues were observed for for MMF + CIC (78.0% [95%CrI: 38.3; 93.6] at 12 months and 88.9% [95%CrI: 43.6; 98.8] at 24 months*).*

To increase compatibility, the company decided to substantially reduce the amount of studies included in the estimation of different outcomes by dropping studies initially included in the NMA. Consequently, direct evidence was lost for specific time points thus reducing the overall sample size of the evidence base.

While at the primary review the majority of the studies failed to meet the inclusion criteria, at the secondary review, no studies met the inclusion criteria. Therefore, no relevant cost-effectiveness studies were found. For this reason, a de-novo analysis was undertaken.

5.1.2 Novartis' economic model submitted to NICE

We now turn to the economic evaluation that Novartis presented to NICE. Novartis report costs per QALY estimates for EVR+rTAC with concomitant corticosteroids for the initial 6 months of therapy compared to MMF+TAC and AZA+TAC both with concomitant corticosteroids for the initial 6 months of therapy.

The model was built in Microsoft Excel©. Here, we summarise the main features of the model.

Throughout their submission Novartis acknowledge that *the advantage of competing immunosuppression regimens is not through limiting rejection (as this is managed very well currently) but it is obtained by improvements in side-effects such as impaired renal functioning*. Therefore, the economic analysis undertaken by the company intends to demonstrate the benefit of EVR in terms of its kidney sparing effect, whilst guaranteeing at least the same effectiveness as the comparator drugs in terms of graft survival.

- #### Model structure

Novartis' cost-effectiveness model was developed as a patient simulation model. The structure of the economic model[1], illustrated in Figure 8, includes a core hepatic model and a renal sub-model and is reported to be appropriate and reflective of the clinical pathway of immunosuppression therapy after liver transplantation.

Novartis state that the use of a patient simulation model is in line with the DSU technical support document 15 (Davis et al, 2014) recommendations as the technical report suggests this modelling approach is appropriate when the patient flow is determined by time since last event or history of previous event.

Novartis also explain that the decision to capture the renal sparing effect associated with EVR+rTAC (demonstrated in the H2304 trial) through a renal sub-model *was considered important because the treatment effect has an impact on more than one aspect of patients' health.* Additionally it is mentioned that *ISPOR good research practice guidelines encourage the design option of using sub-models to simplify the model structure (Karnon et al., 2012).*

Furthermore, the company reports that a patient level simulation approach facilitates *between-model calculations* involving the hepatic core model and the renal sub-model. However it adds that *in the patient simulation, events that occur in the core* [hepatic] *model do not impact on*

[1] The "economic model" refers to the core hepatic model and the renal sub-model throughout the rest of the document.

estimates of the mean change in eGFR from baseline to 12 months, for the different treatment regimens.

The transition probabilities in the renal sub-model define the progression of CKD for each patient. The patient stating level of CKD is randomly generated through a simulation thus it is not possible to report the transition probabilities. However, the decrease in renal function (dependant on the treatment regimen) applied to the baseline CKD stage is fixed and is reported in Table 23. So for example, if a patient is in the EVR+rTAC arm, they will see a reduction of 23.1 mL/min/1.73 m^2 in their eGFR levels, which will correspond to a specific stage of CKD, depending on the patient starting level.

Novartis didn't provide details as to how the NMA eGFR data (originally presented as the difference in eGFR change from baseline) were used to derive the eGFR decrease at 12 months for the different treatment regimens used in the economic model, presented in Table 23.

Upon the ERG's request for further detail on how the company used the NMA data to obtain the estimates in Table 18, Novartis provided the tables shown in Table 24 and reported that these are the absolute results for the outcomes presented in the NMA (corresponding to tables 23-37 in the submission).

The company added that the values reported in **Error! Reference source not found.** show that AZA+CIC is the more effective treatment in terms of preserving renal function at 12 months, followed by EVR+rTAC, MMF+rTAC and finally, standard TAC. It was also stated that the credible intervals show some overlap between therapies.

Although unclear, it seems that, the estimates for the decrease in eGFR at 12 months used in the comparator arms of the economic model (**Error! Reference source not found.**) are based on the one reported for standard TAC in the NMA (**Error! Reference source not found.**).Novartis justify this by arguing that as ciclosporin has limited use in the UK (approximately 5% of current market - Novartis data on file 2014), and the comparators of interest are used in combination with standard dose tacrolimus, only the absolute value for standard dose tacrolimus was used in the economic model.

The minimum method assumes that the minimum utility value for simultaneously occurring health states is considered, for example, the utility for CKD stage 5 is 0.4 and utility for AR is 0.7. The joint state utility for those with both CKD stage 5 and acute rejection is 0.4.The additive model would produce instead a utility of 0.1 (=1-[1-0.4+1-0.7]). The multiplicative model instead would yield 0.28 (0.4 X 0.7).

More details on the other methods can be found on the company's submission page 177.

The company excluded from the analysis health effects associated with hyperacute rejection, HCV and HCC as it was considered that the choice of immunosuppressive regimen has no impact on these outcomes (thus there would not be differences between treatment arms).

- ### Adverse events

Novartis estimated the occurrence of AE associated with the different treatment regimens in the economic model based on NMA data and standard product characteristics (SPCs) for the respective drugs' respective SPCs.

Besides the impact on renal functioning, other treatment related AEs considered were hypertension, diabetes mellitus, infections (bacterial, opportunistic, cytomegalovirus - CMV and fungal), tremor and insomnia.

The recurrence of HCV and HCC was not considered in the model as Novartis state that clinical literature and consultation with clinical experts indicated that there is no effect of the immunosuppressive regimen on these variables.

AE data are reproduced in Table for the different treatment arms. Table 30 and Table 31 report the disutility values and the costs associated with the different AEs, respectively. Costs were reported to have been inflated to 2013 prices where necessary.

The systematic literature searches undertaken by the company did not identify any relevant QoL data associated with specific immunosuppressive-related AEs. Therefore the company used Sullivan et al, 2011 to obtain QoL data.

Table 29. Incidence of AEs in the economic model

AE	EVR+rTAC	AZA+TAC	MMF+TAC
Hypertension	40.3%	57.8%	23.9%
New onset diabetes	15.7%	18.3%	11.6%
Infection	65.7%	60.3%	62.6%
Herpes	0.4%	5.9%	10.1%
Tremor	10.2%	35.5%	33.9%

Table 36. Critical appraisal checklist from Drummond and colleagues (Drummond et al. 1997)

Item	Critical appraisal	Reviewer comment
Has the correct patient group/population of interest been clearly stated?	?	There are some differences between the trial population and the typically presenting UK population. Furthermore there is considerable heterogeneity across study populations in the NMA.
Is the correct comparator used?	?	The ERG believe that reduced TAC monotherapy should have been included as a comparator. since this is widely used in clinical practice; Furthermore, AZA and MMF given with standard doses of TAC do not necessarily reflect clinical practice at later stages of immunosuppressive therapy. These drugs should have been considered with concomitant doses of reduced TAC.
Is the study type reasonable?	?	A patient simulation model was used. The ERG are not convinced this is necessary.
Is the perspective of the analysis clearly stated?	✓	UK NHS PSS
Is the perspective employed appropriate?	✓	NHS Reference Costs
Is the effectiveness of the intervention established?	?	Quality of H2304 is good in establishing the renal sparing effect of EVR+rTAC. The effectiveness absolute values obtained for the NMA rely on a non-robust analysis.
Has a lifetime horizon been used for analysis, if not has a shorter time horizon been justified?	X	The model ran for 80 years. The time frame seems unnecessarily high. After 80 years in the model, these patients will be, on average, 134 year old.
Are the costs and consequences consistent with the perspective employed?	✓	All costs are presented from the UK NHS & PSS perspective
Is differential timing considered?	✓	All future costs and benefits are discounted with a 3.5% rate.
Is incremental analysis performed?	✓	
Is sensitivity analysis undertaken and presented clearly?	?	Probabilistic sensitivity analyses is reported but it lacks robustness.

Note: ✓ indicates ‘clear’; X indicates ‘concerns’; ? indicates ‘some concerns’.

It is worth specifying that the model developed by the company is a patient simulation state-transition model (as opposed to, for example a discrete event simulation model, which also falls under the broader characterization of a patient simulation model).

State-transition models consist of a discrete set of mutually exclusive health states which are evaluated at regular intervals (model cycles). When applied at the patient level (i.e. within a patient simulation framework) these are evaluated stochastically using samples drawn from statistical distributions to determine whether an individual patient experiences a particular transition given the probability of that transition occurring in that particularly cycle.

It is the ERG opinion that Novartis did not provide enough evidence to justify their methodological approach, nor did the company provide clear details regarding the approach and assumptions used.

Novartis state that the use of a patient simulation model is appropriate when the patient flow is determined by time since last event or history of previous event (David et al, 2014). However there is no clear explanation as to why this argument applies to their analysis. Time dependency in the model only exists in relation to transplantation, which happens 30 days before the model begins. This means that this is a “fixed event” in time within the model timeline.

In fact, the patient flow in the economic model is only affected by time since last event in the severe chronic (SCR) rejection state. In the renal sub-model transition probabilities are not time dependant and also do not depend on history of previous events.

The ERG are not convinced that a patient simulation model is necessary to deal with the time dependency of transition probabilities in terms of occupation of subsequent health states in this case as:

- The change in transition probabilities over time from the SPT to the AR state (Table 22**Error! Reference source not found.**) is only dependant on time since transplantation. As time since transplantation progresses simultaneously with the succession of cycles in the model, the change in the transition matrix would be straightforward to implement in a cohort state-transition model. Furthermore, the model considers 3-month cycles and after 12 months all transition probabilities are assumed constant in the model. The same argument applies to transition to the MCR state.

- Episodes of AR are assumed to occur with the same probability, regardless of the number of AR episodes previously experienced by the patient, making the previous history of AR events irrelevant.

- Patients are assumed to stay in the SCR state for 2 cycles as they are assumed to be re-transplanted only after the end of six months. Mortality rates in the SCR state also change from the first cycle spent in the state to the second one. However, this would be easily resolved by creating a tunnel state in a traditional cohort approach. Dividing the SCR state into SCR1 and

- Patient-level state-transition models built in excel require great care as the logic required to specify the transition matrices may be complex and difficult to check for errors (Davis e al, 2014).

- PSA is more difficult to run. Due to its computational burden, Novartis ran 5,000 simulations in their PSA (compared with the 10,000 simulations used in the base case). This is a problem as it affects the reliability of results (this is further explored in the next section).

- It is considerably more difficulty to understand how sensitive the model outcomes are to changes in the model parameters or assumptions. On one hand there is a greater computational time requirement (it took the ERG around 1 hour to run the model every time we changed an input value) and on the other hand given that the model outcomes are generated by randomly sampling parameters it is not possible to replicate the random number streams and thus isolate the impact of changing one particular aspect of the model.

2 Core hepatic model and renal sub-model

Novartis explain that the decision to capture the renal sparing effect associated with EVR+rTAC through a renal sub-model *was considered important because the treatment effect has an impact on more than one aspect of patients' health.* Additionally it is mentioned that *ISPOR good research practice guidelines encourage the design option of using sub-models to simplify the model structure (Karnon et al., 2012).*

Firstly, the ISPOR guidance cited by the company refers to discrete event simulation (DES) models. These are complex models, and even though they fall under the patient simulation model umbrella, cannot be compared with the company's patient simulation state-transition model from a methodological and technical complexity point of view. Therefore, the argument that sub-models simplify the model structure doesn't really apply to the company's model.

Secondly, and more importantly, the company reports that declining renal function can have a major impact on graft and patient survival, and the predisposition to cardiovascular events, as well as result in an increased risk of hospitalisation, hepatic allograft dysfunction and mortality.

Thirdly, Novartis state several times that while the core hepatic model attempts to track patients post-transplantation, the focus of this submission is to demonstrate the renal sparing effect of EVR+rTAC. This is in line with the H2304 clinical trial, where the statistical hypothesis under study were to test the non-inferiority of EVR+rTAC compared to TAC in terms of preventing acute rejection and organ loss and to test the superiority of EVR+rTAC compared to TAC in terms of preserving renal functioning.

Therefore the ERG believe that more emphasis should have been placed on the renal component of the economic model but also that more interaction between the 2 models should have been

considered (perhaps within one broader model structure) as it is known that immunosuppressive therapy for liver post-transplantation has an impact not only on renal functioning but also, equally important, that renal functioning has an impact on graft survival.

3 Specific issues identified in the model

The ERG find that the reporting of the model structure and its assumptions lacked clarity. Also few justifications were provided as to why those assumptions were considered necessary and appropriate.

In this subsection we discuss the different health states included in the economic model, the errors in formulae used to allocate patients into different health states, the cycle length of the model, the time horizon used and finally the number of simulated patients.

Health states in the economic model

The company's model structure is presented in Figure 8 above. The ERG found some inconsistencies in the representation of the model structure and the description of health states.

The hepatic core model describes 6 health states:

Stable post-transplant state (SPT):

Even though the company claims that patients enter the model in the SPT state 4 weeks after transplantation, this is not entirely accurate. Patients in the comparator arms of the model are assumed to enter the model immediately after transplantation, and incur the respective costs.

During the clarification process, the ERG asked the company to clarify at what point after transplantation the model begins. The company clarified that the model starts immediately (one day) post-transplant, as the use of comparator therapies begins at this time point. However, for the EVR+rTAC arm of the model, patients only start receiving treatment 30 days after entering the model (i.e. 30 days post-transplantation). Costs and benefits for the EVR+rTAC arm were adjusted to reflect the later starting point accounting for 2 of the 3 months in the first cycle. The ERG found a mistake with regards to the AEs considered in the first cycle of the model. This s explored in the AEs subsection.

Acute rejection (AR):

Patients in the SPT state, suffering from AR (diagnosed by biopsy) move to this state, where they are given a high-dose steroid course of treatment for a minimum of 3 days and a maximum of 2 courses (i.e. 6 days of steroid treatment). If the treatment is successful, patients return to the stable SPT state. If treatment fails then patients move the ASRR state.

Of note is that the company mentions a maximum of 3 courses of steroid treatment. Nonetheless, a maximum of 2 courses was considered in the base case economic analysis.

Acute steroid resistant rejection (ASRR):

If steroid treatment fails, patients move to the ASRR state. In this state patients will undergo anti-lymphocyte therapy for 14 days. If the treatment is successful then patients return to the SPT state. If treatment is unsuccessful, patients move to the SCR state. Despite the fact that arrows in Figure 8 seem to indicate that patients can remain in this state for longer than 1 cycle, this is not the case as per the transition probability matrix used in the excel model.

Severe chronic rejection (SCR):

If patients fail both steroid and anti-lymphocyte treatments for acute rejection, they will move into the SCR state. It is assumed that all patients entering the SCR state will suffer graft failure and will require a re-transplant.

Patients are assumed to stay in the SCR state for 2 cycles as they are assumed to be re-transplanted only after the end of six months (so only after being 2 cycles in this state).

Under the current structure, re-transplant patients go straight back to the SPT state, but EVR patients are only meant to enter this state 4 weeks after having their transplant. Therefore the model is not capturing this 4-week period during which events like hyperacute rejection and any other early rejections may happen. This could potentially bias the results as the outcomes captured for re-transplantation patients are better than they would be in real life. Nonetheless, this doesn't impact the economic base case analysis as virtually no patients are re-transplanted in the model across all treatment arms. Furthermore, as EVR is not expected to have an impact on the prevention of hyperacute rejection, the marginal difference across model arms could potentially even out, provided AZA and MMF also do not have an impact on this outcome.

Mild chronic rejection (MCR):

Patients are reported to be able to develop MCR from any state except SCR and death. However, that is not the case, as in the excel model patients are only allowed to transition to the MCR state from the STP state. This is an asymptomatic state and it has been assumed that patients can only move to this state 1 year post-transplant. Moreover, due to some issues found in the excel formulae, patients can only progress to this state after they have completed 1 year in the STP state. This is discussed in the next section of the report.

The company also mention that patients can go from this state to the STP (Table 45 in the submission), however that is not the case in the base case economic analysis, where patients reaching the MCR state don't leave the state unless dead. This means that patients can stay in this state for years.

Clinical opinion sought by the ERG revealed that the state of mild chronic rejection seems somewhat vague and unfamiliar. Our clinical expert advisor did not see any valid (or justifiable) reason for patients only to progress to such state 1 year after transplantation. In fact the company recognizes that it is not clear what causes this state and that it is an asymptomatic one. The ERG question the clinical plausibility of this state and the plausibility of patients remaining in the state for years. The relevance of including this health state in the model is not clear. There is no change in the utility of life of patients when in this state, however the resource use associated with MCR is similar to that of the AR state. This is further explored in the next section.

Hepatic-graft related death (HD):

Patients were assumed to have a higher mortality rate in the SCR state of the model. As patients are virtually kept on this state for 2 cycles, a 3% mortality rate is used during the 3 initial months of the patient being in the SCR state to reflect the mortality rate of patients dying while on the waiting list. After the 6 months, when the patient will undergo surgery, a 0.7% operative death rate is applied to reflect patients dying during surgery.

The ERG found some mistakes in the formulae related to the mortality rates. This will be further explored in Section 5.2.3.

The health states captured in the renal model were taken from the NICE guidance on CKD progression and are presented inTable 21. The model structure seems to accurately capture the evolution of CKD disease. However in the context of the decision problem faced in this submission, the ERG's opinion is that a more interactive model linking hepatic and renal outcomes would be better suited.

Table 25 shows the increase in mortality associated with the different stages of CKD. It is appropriate to account for the increase in mortality as renal disease progresses.

A background natural mortality rate was also used for both hepatic and renal models, which is appropriate.

Allocation of patients to health states in the model

In a patient simulation state-transition model, for each cycle the probability of each transition probability is compared with a random number to determine if that particular transition occurs in that cycle. When the random number is less than the probability, the transition element is set to 1 thus indicating that the transition occurs. If the random number is higher than the probability then that specific transition does not occur.

The ERG found a mistake in the formulae used to allocate patients to the different health states in the model.

To model the 3 health states for which transition probabilities are dependent on time since transplantation (STP, AR and MCR) the formulae need to consider time elapsed since

of the possible transitions is acted on (Davis et al, 2014). Novartis have assumed that deaths have precedence over other events.

According to David et al,2014 this can create a bias towards the more extreme event which ensures that rarer but potentially more dangerous events are not ignored within the model but this means that the model is likely to favour treatments which prevent the more severe events over those that prevent the less severe events. The bias that this generates within the cost-effectiveness estimates can be minimised by reducing the cycle length, as this lowers the probability that two events are sampled to occur within a single cycle.

Cycle length

The cycle length in the economic model is 3 months and a half-cycle correction was not applied.

The ERG are concerned that 3-month cycles could not capture all the relevant outcomes for the disease modelled as it seems that disease progression is faster in real life than in the model. The company submission stated that for example, if patients failed the first steroid treatment in the AR state, this would be repeated 3 days later with the same success probability. Therefore, the ERG have raised this as a question of concern during the clarification stage.

The company explained that *the 3-month cycle length was validated with clinical experts who advised that the timing of any acute rejection was likely to occur within the three month period, but that the majority of patients would move back to the stable state after treatment*.

Novartis added that clinical opinion was that AR and ASRR are mainly asymptomatic stages, *where no additional healthcare costs would be incurred or benefits be gained* compared with the stable state, thus they would not expect that shortening the cycle length would have an impact on the results of the model, *as the costs of the acute rejection state are the same as the SPT state (with the exception of the relatively small cost of high dose corticosteroids) and the utility is also the same.*

The clinical opinion provided by the ERG's clinical advisor is that 3-month cycles seem too long to capture all the relevant events and that monthly cycles would perhaps be more appropriate. Furthermore the ERG find Novartis' argument that costs and benefits associated with the AR and ASRR are similar to those in the SPT state quite surprising as this is not the case. Not only does the submission report very different resource use for these states but this is also the case for the economic model, where ASRR costs (£6,506) are ninety-fold the SPT ones (£72) and AR costs (£1,922) are twenty-seven-fold the SPT ones.

Time horizon

The time horizon considered in the economic model was 320 cycles (80 years). Given that the average starting age of patients is 54 years (with a standard deviation of 10 years to allow for random sampling of this parameter), the time frame seems unnecessarily high, adding to the computational burden of the model. After 80 years in the model these patients will be, on average,134 year old. After 40 years in the model (i.e when patients are on average 94 years), 100% of patients are dead.

Number of simulations

The model simulated 10 000 individual patients. The ERG didn't find any justification for the number of patients selected, or any mention to this parameter throughout the submission. Methods of justification can include a graphical representation of the costs, QALYs and the cost per QALY gained and determining at what number of patients the estimated standard error in the results appear acceptable (Davis et al, 2014).

If the results are found to vary significantly when selecting a different random number stream then the model should be checked to see whether there are any unintended correlations between samples that are supposed to vary independently.

The ERG are concerned with the number of simulations and the lack of stability in the patient simulation model. The ERG have run the Novartis simulation model for the base case as submitted by the company, **without any changes,** to test the model stability with regards to ICER results. After running the model two times, the ERG found a considerable variation in the ICERs reported (Table 41), especially in the case of the EVR+rTAC versus MMF+TAC ICERs, which varied between £110,797 and £120,651 (nearly a 9% change).

To also note is that the variation observed in the ICER for EVR+rTAC versus AZA+TAC is smaller than the one observed in the ICER for EVR+rTAC versus MMF+TAC arm. Nevertheless it is still a significant one.

The ERG tried to run the model with 15,000 simulations but not only did the company made this option not possible by default in the model (the cell input value for the number of simulations was set to allow a maximum of 10,000) but also, when the ERG changed this definition and ran the model with 15,000 simulations the model break near to the end of the simulation process.

calculation of transition probabilities within the model. Additionally the model used several parameters taken from available literature.

In this subsection we focus on the model parameters used in the hepatic and the renal models and describe how transition probabilities between health states were estimated within the different arms of the main economic model.

The ERG have several concerns with the clinical parameters used in the economic model.

H2304 study – TAC trough levels

The main source for the effectiveness of EVR+rTAC is H2304 trial. Novartis acknowledge that one of the limitations of the trial is the fact that the target ranges of TAC are higher than the ones observed in standard UK clinical practice (where through levels <5 ng/mL are considered as low-dose TAC). This is attributed to the trial being set in the US and changes in clinical practice since 2007.

As previously mentioned, in the EVR+rTAC arm of the trial the planned dose of TAC was to achieve the 3 – 5 ng/mL by 3 weeks after randomization and keep on these for the remainder of the study. In the standard TAC or the trial, the trough levels were targeted to be maintained at 8-12 ng/mL until month 4, and at month 4 TAC whole blood trough levels would be decreased to a target trough level of 6-10 ng/mL for the remainder of the study. Finally, in the TAC elimination arm (i.e. the EVR arm), after EVR blood trough levels were confirmed to be in the target range (3-8 ng/mL), TAC tapering was started, achieving a target TAC trough level of 3-5 ng/mL by 3 weeks after randomization. TAC elimination began at month 4 post-transplant and TAC was completely eliminated by the end of month 4.

In Figure 11 below the ERG present the TAC trough levels reported in the CSR for the 12-month analysis.

It can be observed that in general the average trough levels of TAC were higher than the ones initially planned for all arms of the trial. In fact the “reduced” TAC arm shows trough levels above 5 ng/mL throughout the 12 months. Clinical opinion sought by the ERG indicated that the trough levels observed in the reduced TAC arm are comparable to what would be considered a standard TAC regimen in the UK. This means that the standard TAC arm in the trial is also not reflective of UK clinical practice, presenting extremely high trough levels. Our clinical advisor explained that a standard TAC regimen in the UK is 6 – 8 ng/mL until month 1, just above 6 ng/mL until month 4 and then between 5 and 6 ng/mL until the end of first year.

It is also noticeable that in the TAC elimination arm, the average TAC trough levels after month 4 ranged from [redacted] and that while some patients seemed to be at 0 trough levels, this was not the case for all patients. It is therefore possible to hypothesise that the TAC trough levels in the TAC elimination arm are closer to what would be considered a reduced level of TAC in the UK than any other trial arms in H2304 (at least from month 4 to month 9). This leads to the question:

SPT to MCR (stable post-transplant to mild chronic rejection)

The company reports to assume that this transition can only occur 1 year after transplant, however the ERG found a mistake in the model with regards to the allocation of patients to the health states (discussed in Section 5.2.2 of our report) which means that patients can only move to the MCR state 1 year after being in SPT state (and not 1 year after transplantation).

The company reported this value to be 4% as per clinical opinion. However the value used in the economic model is 1%. The company make no mention to the calculation involved in obtaining this value but the ERG assume this is 4%/4 months = 1% every 3 months. The same issue raised above for dividing transition probabilities by time intervals applies here, even though the final result is virtually the same when using the correct method.

Surprisingly, with 1% of patients moving to this heath state every month, no apparent clinical plausibility to justify its existence and no determent in QoL associated with it, according to the company sensitivity analysis **when the MCR state is eliminated from the analysis, the final ICER comparing EVR+rTAC with AZA+TAC is £176,410, while the ICER for EVR+rTAC compared with MMF+TAC is £233,331**. (Of note is that base case ICERs are £187,842 for EVR+rTAC vs AZA+TAC and £110,797 for EVR+rTAC vs MMF+TAC). This is further explored in Section 6.

AR to SPT (acute rejection to stable post-transplant)

The company reports using the 86.3% probability of successful first line steroid therapy and source the value back to Aydogan et al, 2010. However, in the economic model the probability used to account for patients going from AR to SPT is 98.1%. Novartis arrive to this estimate by undertaking the following calculation:

1 – [(1 – 86.3%) ^ (number of steroid treatment courses)]. As they assume number of steroid treatment courses = 2, then:

1 – [(1 – 86.3%) ^ (2)] = 98.1%.

From a conceptual point of view, the formula can be translated as:

> 1 – [(probability of failing steroid treatment) * (probability of failing steroid treatment)] which is equivalent to the probability of success of 1st line steroid treatment + probability of 1st line steroid treatment failing and second line steroid treatment being successful.

However, two things should be considered:

- The 86.3% of successful steroid treatment in Aydogan et al, 2010 is for a 1 course (3 days) treatment. Nonetheless Novartis assume the same probability of successful resolution of AR for 1st and subsequent episodes.

ciclosporin with reduced TAC. The same is applicable to the McDiarmid et al, 1993 study, where TAC trough levels were maintained below 2 ng/mL.

Overall, the ERG question the validity of the NMA results for the renal outcomes. The allocation of different studies' treatment arms to the reduced and standard TAC categories is inconsistent and misleading.

As the standard TAC connector across the NMA studies is so heterogeneous (note for example that in the Boudjema et al, 2009 study the trough level of TAC in the standard TAC arm was close to 12 ng/mL for several months during the first year) then the NMA results are likely to lack robustness.

Regardless of the ERG concerns with the validity of these estimates, looking atTable 24, it can be argued that the estimate for the decrease in renal function at 12 months for the MMF+ "reduced" TAC arm (28.2) should be used in the MMF arm of the economic model instead of the standard TAC estimate (31.6). This is because the MMF+ "reduced" TAC is actually closer to a MMF+ standard TAC arm if we take the H2304 study (and UK clinical practice) as reference. The same value could be used for the AZA arm of the model as the renal dysfunction is determined by the levels of TAC (and is not dependent on the concomitant drug).

Table 43. TAC trough levels in the NMA studies.

H2304	Standard TAC	The standard TAC arms presents average trough levels of 10 ng/mL at randomization, around 9 ng/mL until month 9 and then close to 8 by the end of the 1st year.
	EVR + reduced TAC	The reduced TAC arm presents average trough levels of 11 ng/mL at randomization and then around 6 ng/mL from month 4 until end of the 1st year.
(Neuberger *et al.*, 2009[65])	Standard TAC	The standard TAC arm presents average trough levels of 11 ng/mL until month 3 and then decreased to 9 by the end of the 1st year.
	MMF + reduced TAC	The reduced TAC arm presents average trough levels of 9 ng/mL until month 3 and then decreased to 8 by the end of the 1st year.
(Boudjema *et al.*, 2011[83])	Standard TAC	The standard TAC arm presents average trough levels above 10 ng/mL for the whole of the 1st year.
	MMF + reduced TAC	The reduced TAC arm presents average trough levels between 10 and 9 ng/mL for the 1st year.
(Porayko *et al.*, 1994[88])	Standard TAC	TAC trough levels were maintained below 1.0 ng/mL at month 12, whilst being reported to be higher in the first 4 weeks- with only limited data being reported.
	AZA+ciclosporin	
(McDiarmid *et al.*, 1993[93])	Standard TAC	TAC trough levels were maintained below 2 ng/mL.

Table 48. Post-transplantation ongoing care costs in the renal model

Health state	Time	Cost accrued in the model
CKD stage 5	14 cycles (4 years)	CKD stage 5 costs
No CKD (after transplantation)	52 cycles (13 years)	Transplantation cost (one off) + Post-transplant ongoing care costs
CKD stage 1/2	19 cycles (5 years)	CKD stage 1/2 costs + Post-transplant ongoing care costs
Death	Remaining cycles	No costs

Finally, all of the costs presented for the CKD health states were wrong in the submission (when compared with the economic model). This was is a misreporting problem, as the resource use and the unit costs in the submission are correct and match the ones in the excel model, used to derive the total costs.

Intervention and comparators' costs

Novartis have used the Prograf brand price to cost the use of TAC in the model, however there is no justification given for this decision. The ERG looked at the prices for different brands of TAC, and Prograf is the most expensive one (£1.61 per mg), with Vivadex being the least expensive (£1.2 per mg). As there is no apparent justification for choosing Prograf over any other brand used in the UK, the ERG took the average across all brand prices and derived an estimate of £1.3 per mg. Ideally the company would have used a weighted average, according to market shares.

No on-going monitoring costs were included as it was assumed that that these were included in the consultations costs (page 191 of Novartis submission)

The company states that dose assumptions for EVR+rTAC were taken from H2304, while dosages for other drugs were based on the BNF. These are presented in Table 49. The ERG could not find the TAC doses used in H2304 in the CSRs, as TAC prescription is usually reported as trough levels to be achieved and not dose. This was also the case for most studies in the NMA, where TAC was prescribed according to target trough levels.

Clinical opinion sought by the ERG advised that even though it is reasonable to assume that drug doses will remain overall constant after 1 year, the primary focus is to keep TAC doses as low as possible for a specific patient, hence some of these doses were considered slightly high to be maintained for the rest of the patients' lives. Although the MMF SPC recommends a dose of 3000mg per day, the ERG's clinical advisor also mentioned that the doses of oral MMF seem high after 1 year, and that 1000 or 2000 (and not 3000) mg would reflect UK clinical practice more accurately for the 1 year post transplantation period.

For scenario 4, Novartis decreased the baseline eGFR level from 81 mL/ min per 1.73 m^2 to 60mL/min per 1.73 m^2. The reason provided for this was that the eGFR baseline levels in H2304 might be higher than the ones usually observed in UK clinical practice.

Similarly to the other scenario analysis, there was not consistency in the impact on the final ICERs across treatment arms. Compared with the base case ICERs, the ICER for EVR+rTAC versus AZA+TAC decreased whilst the ICER for EVR+rTAC versus MMF+TAC increased.

Table 54. Scenario 4 run by Novartis

Cost-effectiveness results per patient	EVR+rTAC (1)	AZA+TAC (2)	MMF +TAC (3)	Incremental value (1-2)[4]	Incremental value (1-3)
Total costs £	160,845	131,392	129,448	**£29,453**	**£31,397**
QALYs	3.79	3.63	3.62	**0.16**	**0.17**
ICER				**£184,081**	**£184,372**

5.3.2 Probabilistic sensitivity analysis results

Novartis provided the results of a PSA using 1,000 simulations of 1,000 patients. According to the company, the results were stable at this level of simulation. The company tested a PSA using 1,000 simulations of 5,000 patients and claimed to have similar results. Therefore as the model run time was long, the 1,000 simulations of 1,000 patients were chosen as a practicable solution, according to the company.

As previously mentioned by the ERG, it is likely that the base case simulation model is not stable with 10,000 patients (i.e. simulations). Therefore, running the model with 1,000 is even more likely to generate unreliable estimates.

The results for EVR+rTAC versus AZA+TAC showed an incremental cost per QALY gained of £184,714 (compared to the base case ICER of £187,842).

The results for EVR+rTAC versus MMF+TAC showed an incremental cost per QALY gained of £105,526 (compared with the base case ICER of £110,797).

Again, the ERG note that these results are based on a small number of simulations and are likely to lack robustness.