in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Axitinib for the treatment of advanced renal cell carcinoma after failure of prior systematic treatment

Produced by Kleijnen Systematic Reviews Ltd in collaboration with Erasmus University Rotterdam and Maastricht University

Authors
Rob Riemsma, Reviews Manager, KSR Ltd
Maiwenn Al, Health Economics Researcher, EUR
Isaac Corro Ramos, Health Economics Researcher, EUR
Sohan Deshpande, Systematic Reviewer, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Caro Noake, Information Specialist, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Johan L Severens, Professor of Evaluation in Health Care, EUR
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to
Rob Riemsma, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 16/10/2012

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 11/87/01 STA.

Declared competing interests of the authors
None.

Acknowledgements
The authors are grateful to Saskia de Groot (EUR) for her input in the clarification phase.

Commercial in confidence data are redacted throughout the report.

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

This report should be referenced as follows:
Riemsma R, Al M, Corro Ramos I, Deshpande S, Worthy G, Noake C, Armstrong N, Severens JL, Kleijnen J. Axitinib for the treatment of advanced renal cell carcinoma after failure of prior systematic treatment: a Single Technology Appraisal. York: Kleijnen Systematic Reviews Ltd, 2012.

Contributions of authors
Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Maiwenn Al acted as health economic project lead, critiqued the manufacturer's economic evaluation and contributed to the writing of the report. Isaac Corro Ramos and Nigel Armstrong acted as health economists on this assessment, critiqued the manufacturer's economic evaluation and contributed to the writing of the report. Sohan Deshpande acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the manufacturer's submission and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Johan L Severens critiqued the manufacturer's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the manufacturer's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AE	Adverse Events
AJCC	American Joint Committee on Cancer
bd/b.i.d	Twice Daily
CE	Cost Effectiveness
CEA	Cost-effectiveness Analysis
CEAC	Cost effectiveness Acceptability Curve
CI	Confidence Interval
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
ECOG PS	Easter Cooperative Oncology Group Performance Status
EMEA	European Medicines Agency
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
EUR	Erasmus University Rotterdam
FKSI-DRS	Functional Assessment of Cancer Therapy Kidney Symptom Index -Disease Related Symptoms
HR	Hazard ratio
HRQL	Health-related Quality of Life
HTA	Health Technology Assessment
IC	Indirect Comparison
ICER	Incremental Cost-effectiveness Ratio
ITT	Intention to Treat
KSR	Kleijnen Systematic Reviews
LY(S)	Life Year (Saved)
mg	Milligram
MS	Manufacturer's Submission
MSKCC	Memorial Sloan-Kettering Cancer Centre
MTC	Mixed Treatment Comparison
NHS	National Health Services
NICE	National Institute for Health and Clinical Excellence
NIHR	National Institute for Health Research
NR	Not Reported
od	Once Daily
OR	Odds Ratio
OS	Overall Survival
ORR	Objective response rate
PCT	Primary Care Trust
PD	Progressive Disease
PDGFR	Platelet-derived growth factor receptor
PF	Progression Free
PFS	Progression Free Survival
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services
QALY(s)	Quality-adjusted Life Year(s)
RCC	Renal Cell Carcinoma

RCT	Randomised Controlled Trial
RECIST	Response Evaluation Criteria in Solid Tumours
RPSFT	Rank preserving structural failure time
RR	Relative Risk
SAE	Serious Adverse Events
SD	Standard Deviation
SE	Standard Error
STA	Single Technology Appraisal
TNM	Tumour Node Metastases system
UMC	University Medical Centre
VEGFR	Vascular endothelial growth factor receptor

Table of Contents

Table of Tables

Table of Figures

1. SUMMARY

1.1 Scope of the manufacturer submission

This is the summary of the ERG report on the manufacturer's submission: Single Technology Appraisal (STA) for "Axitinib (Inlyta, Pfizer) for the treatment of advanced renal cell carcinoma after failure of prior systemic treatment". The objective of this STA as defined by the final scope is:

> *To appraise the clinical and cost effectiveness of axitinib within its licensed indication for the treatment of advanced renal cell carcinoma after failure of prior systemic treatment.*[1]

The scope of the manufacturer's submission is consistent with the components of the question and approach outlined in NICE's final scope. The anticipated authorised use of axitinib, an oral drug, is for the treatment of adult patients with advanced renal cell carcinoma (RCC) after failure of prior treatment with sunitinib or a cytokine. Marketing authorisation from the European Commission is expected in September or October 2012 *(Source: Pfizer's submission, Section 1.3 and 1.5, page 17-18*[2]*).*

The only comparator specified in the scope is best supportive care (BSC).

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

There are two populations of interest: cytokine refractory patients and sunitinib refractory patients.

For cytokine refractory patients, the comparison between axitinib and BSC relies on an indirect comparison using evidence from two trials: AXIS (axitinib versus sorafenib) and TARGET (sorafenib versus placebo).

For sunitinib refractory patients, the comparison between axitinib and BSC relies on a simulated treatment comparison (STC). STC is a novel technique to derive indirect comparisons between competing treatments (say A and B). Unlike mixed treatment comparisons (MTCs) which provide an average measure of the difference between A and B across all studies, STCs aim to answer a more specific question: what difference could we expect if A and B had been compared in the same trial (*Source: Pfizer's submission, Section 6.7.2, page 80*).

In this submission the STC relies on individual patient data (IPD) for the axitinib-arm from the AXIS-trial, and summary data for the placebo-arm from the RECORD-1 trial (everolimus versus placebo). This was supported by additional analyses using overall survival data from patients receiving first-line sunitinib followed by sorafenib or BSC from the Swedish Renal Comparison database (RENCOMP).

Effectiveness of axitinib in the AXIS trial:

As can be seen from Table 1.1 the results in terms of overall survival (OS), progression-free survival (PFS) and response are more favourable for axitinib in the cytokine refractory population than in the sunitinib refractory population. Nevertheless, PFS shows a significant result in both populations, while OS shows no significant results in either population.

The main outcomes from the AXIS trial are as follows (see Table 1.1):

Table 1.1:Main outcomes from the AXIS trial (axitinib versus sorafenib)

	All patients (N=361/362)	Cytokine refr. (N=126/125)	Sunitinib refr. (N=194/195)
Overall Survival	HR=0.969 (0.800, 1.174)[1]	HR=0.813 (0.555, 1.191)	HR=0.997 (0.782, 1.270)
PFS	HR=0.665 (0.544, 0.812)[1]	HR=0.464 (0.318, 0.676)[2]	HR=0.741 (0.573, 0.958)[2]
Response	RR=2.056 (1.408, 3.003)[3]	RR=2.392 (1.434, 3.992)[3]	RR=1.477 (0.792, 2.754)[3]
HRQoL (EQ-5D)	NR	Post-Treatment difference: [redacted]	Post-Treatment difference: [redacted]
AEs (grade 3-4) - Death (all) - Death due to AE - Any AE - Diarrhoea - Hypertension - Fatigue - PPE	(N=359/355)[4,5] 31.5% vs 30.7% 9.5% vs 6.8% 48.2% vs 52.2% 10.0% vs 7.1% 15.6% vs 11.0% 9.8% vs 3.7% 5.0% vs 16.1%	(N=126/123)[4,6] NR NR [redacted]	(N=126/123)[4,6] NR NR [redacted]

AE=Adverse events, NR=Not reported, PPE=Palmar-plantar erythrodysaesthesia syndrome
[1]) Adjusted for ECOG PS and prior treatment regimen; [2]) Adjusted for ECOG PS; [3]) Risk Ratio for overall confirmed ORR (CR and PR); [4]) Safety population: All patients who received at least one dose of study medication. [5]) Results as reported in the MS. [6]) Results as reported in the response to the clarification letter.

Axitinib versus BSC in cytokine refractory patients:

- For overall survival the Hazard Ratio (HR) for axitinib versus placebo was 0.63 (95% CrI 0.41-0.99). This is based on the ITT population, censored for cross-over in the TARGET trial
- For PFS, the HR for axitinib versus placebo was 0.251 (95% CrI 0.165-0.379)
- Other outcomes, such as response, quality of life and adverse events were not reported.

Axitinib versus BSC in sunitinib refractory patients:

- For overall survival using the ITT RECORD-1 placebo cohort adjusted for cross-over using the RPSFT method, the estimated median OS was 36 weeks (8.3 months) for axitinib-like patients assuming that they received placebo.
- The estimated OS HR for axitinib versus BSC was 0.619 (95% CI 0.384-0.997) using data from a Swedish patient registry.
- For PFS using the ITT RECORD-1 placebo cohort, the estimated median PFS was 6.9 weeks (1.6 months) for axitinib-like patients if they had received placebo.
- Other outcomes, such as response, quality of life and adverse events were not reported.

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

There is no direct evidence for the comparison axitinib versus BSC. Instead the submission relies on an indirect comparison for cytokine refractory population, using evidence from two RCTS and a simulated treatment comparison for the sunitinib refractory population, using evidence from single treatment arms from two trials.

For the indirect comparison in the cytokine refractory population comparison of trial populations used in the analyses is not possible, because patient characteristics are not reported for cytokine refractory patients separately in both trials. Another problem for this comparison is concern about the results for overall survival in the TARGET trial due to treatment switching from placebo to sorafenib at the point of disease progression.

For the sunitinib refractory population, the evidence relies on a simulated treatment comparison, this comparison is not based on randomised treatment allocation, but on a comparison of two single

treatment arms; therefore there is considerable potential for bias in the outcomes of this analysis. In addition, the uncertainty around the STC results is unclear. There is no presentation of an associated SE or 95% CI for any of the results. This is a major omission as we have no idea of the uncertainty of the estimates. The estimates of mean or median PFS or OS and the associated difference between axitinib and placebo should all have been reported with associated variance estimates.

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

The ERG confirmed that there was no existing estimation of cost-effectiveness of axitinib versus best supportive care, and that it was appropriate for the manufacturer's submission to focus on a new cost-effectiveness model.

This was a Markov state-transition cost-utility model implemented in Microsoft Excel which compared treatment with axitinib and BSC with BSC alone, in line with the decision problem. The three health states were progression-free, progressive disease (PD) and death, and the outputs were expressed as cost per quality adjusted life year (QALY). The cycle length of the model was four weeks and the time horizon of the study was 10 years. The analysis was performed for two subgroups, i.e. cytokine refractory patients and sunitinib refractory patients.

The proportion of patients in each health state at each point in time was calculated directly from the parametric survival function equations for both progression-free survival and overall survival. For the axitinib arm, the parametric survival functions are directly estimated from the AXIS trial data. The survival function for BSC arm was estimated using different approaches for the different subgroups, i.e. an indirect treatment comparison for the cytokine refractory group and a simulated treatment comparison for the sunitinib refractory group.

Utilities applied to the health states were based on utility data collected in the AXIS trial using the EuroQoL-5D (EQ-5D) instrument, completed by the study patients at Day one, every four weeks afterwards, at the end of study treatment or withdrawal and at follow up Day 28. Utilities were assumed to be equal for both treatments.

Cost of axitinib were based on the recommended dosing schedule for the product, and adjusted for the relative dosing intensity (RDI) observed in the AXIS trial (102%). No administration costs were included, and treatment costs are associated to the PF state only. In addition, a PAS was approved, which means that all results were calculated using the cost of axitinib with and without PAS. The estimates of routine medical monitoring for the stable and progressed disease states included in the model and the adverse events included (hypertension and diarrhoea) were based on those considered in the PenTAG economic model[3] and the everolimus STA.[4] Health-state costs were applied independent of treatment arm. Costs and quality adjusted life years (QALYs) were discounted at 3.5%. The impact of parameter uncertainty was estimated in a probabilistic sensitivity analysis. Scenario analyses were run on key parameters, especially relating to the extrapolation of the PFS and OS curves for axitinib and the methodology used to model the BSC arm.

The base case ICER (cost per QALY gained) with PAS amount to £65,326 and to [redacted] without PAS for the cytokine refractory subgroup. For the sunitinib refractory subgroup these ICERs are £40,933 with PAS and [redacted] without PAS.

Regarding the scenario analyses, for the cytokine refractory population the manufacturer concluded that the key parameters which increased the ICER included use of a lognormal model to extrapolate

axitinib PFS, and a Gompertz model to extrapolate axitinib OS. None of the scenarios explored by the manufacturer provided an ICER below £50,000.

For the sunitinib refractory subgroup, the manufacturer concluded that the model is robust to the majority of structural assumptions made. Most of the scenarios examined, including extrapolation assumptions, the STC approach, and source of utility estimate, produced ICERs lower than £50,000/QALY (with PAS). The exception was the method of OS comparison used that produced quite different results ranging from values very close to the base case ICER to a maximum difference larger than £20,000. In particular, the use of the RENCOMP model analysis resulted in ICERs higher than the base case (ranging from between £43,384 and £56,113 with PAS). According to the manufacturer, this indicates that the incremental survival benefit assumed over BSC is a key driver of the model result.

The PSA results with PAS showed a 31% probability that the ICER is below £50,000 per QALY for the cytokine refractory population whereas this was 67% for the sunitinib refractory population. Without PAS, these percentages were [redacted] for the cytokine refractory population and [redacted] for the sunitinib refractory population.

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

The economic model described in the MS is considered by the ERG to meet the NICE reference case[5] and is in-line with the decision problem specified in the scope.

The ERG assessment indicated that the model was generally well presented and reported. Some minor issues were identified regarding the use of standard deviations in the PSA instead of standard errors; whilst these errors have a clear impact on the outcomes of the univariate and probabilistic sensitivity analyses, they do not impact the estimate of the central ICER. In addition, the clinical assessment revealed an error in the hazard ratio for the BSC arm in the cytokine refractory subgroup. This error was corrected but this only had a minimal impact on the ICER, which decreased to £64,388 with PAS and [redacted] without PAS.

After rerunning the PSA with the correct standard errors, we found that with PAS there is a 19% probability that the ICER is below £50,000 per QALY for the cytokine refractory population whereas this was 83% for the sunitinib refractory population. Without PAS, these percentages were [redacted] for the cytokine refractory population and [redacted] for the sunitinib refractory population.

Various assumptions around the utility estimates were discussed, and some extra scenario analyses showed that changing these assumptions had very minimal impact of the ICERs.

The manufacturer performed the univariate sensitivity analysis by varying all parameters between plus and minus 20%. This is often not very informative, since this 20% may be either a under- or over-estimate of the true uncertainty. Thus, the ERG performed a univariate sensitivity analysis in which parameters were varied between the limits of their 95% confidence interval (as defined for the PSA). This revealed that for the cytokine refractory subgroup, the ICER is extremely sensitive to changes in the HR for the overall survival. At the upper limit of the 95% confidence interval, the ICER would amount to almost £400,000 (with PAS). This is in sharp contrast with the results for the sunitinib refractory subgroup, where changes to input parameters lead to modest changes in the ICER. This is related to the fact that no measures of uncertainty were provided for the STA adjustment factor for the BSC arm, which means that this uncertainty was not considered in the univariate (or probabilistic) sensitivity analysis.

It is important to realise is that many uncertainties related to the health economic evaluation in the sunitinib refractory subgroup have not been quantified, and thus are not represented in the central estimates of the ICER or in the CEACs. Both the STC approach and the RENCOMP approach to estimating the PFS and OS of the BSC group have the potential for considerable bias, either upwards or downwards. This means that the same is true for the ICERs. In addition, the STC approach lacks an estimation of the uncertainty surrounding the point estimates it provides. Again, this also means that the uncertainty around the sunitinib-refractory ICER is most likely severely underestimated.

For the cytokine refractory subgroup most uncertainties have been taken into account, revealing a large uncertainty in the number of QALYs gained, and thus around the ICER.

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

1.6.1 Strengths

Search methods were clearly presented and reported. The MS provided sufficient detail for the ERG to appraise the searches. Additional searches of conference abstracts and other relevant resources were undertaken by the manufacturer for the clinical effectiveness, cost effectiveness and HRQL sections. The ERG concluded it was unlikely that any of the errors identified in Appendix 1B would have impacted greatly on the recall of searching and that any potentially consequential errors would have been mitigated due to the overlap between strategies and the breadth of resources searched.

Overall, the main strength of the manufacturer submission is a well-constructed and presented case on the clinical and cost-effectiveness of axitinib. The main trial reported in the manufacturer submission is the AXIS trial. This is a high quality randomised phase III open label trial comparing axitinib with sorafenib, in a large purely second-line population. The trial included both relevant populations for this appraisal: a sunitinib refractory (55%, n=389) and a cytokine refractory subgroup (35%, n=251). These subgroups account for about 90% of the ITT trial population.

For the health economic evaluation, the extrapolation of the axitinib data was done thoroughly, with various scenarios explored to assess the robustness of the estimates. In addition, unlike in previous appraisals in mRCC, the current model contained utility values measured (using EQ-5D) alongside the AXIS study.

1.6.2 Weaknesses and areas of uncertainty

In general the searches were constructed in a logical fashion; however, there was some redundant usage of the explosion function, MeSH/Emtree headings and Economics/HRQL filters. CRD guidance recommends against limiting searches to English language papers, as used in the 10.10 Cost Effectiveness and 10.12 HRQL strategies. Without rerunning searches the ERG is unable to say whether this would have excluded any potentially useful non-English language papers.

The main weakness of the manufacturer's submission is that there is no direct evidence for the comparison axitinib versus BSC. Instead, the submission relies on an indirect comparison for cytokine refractory population, using evidence from two RCTs and a simulated treatment comparison for the sunitinib refractory population, using evidence from single treatment arms from two trials.

For the **cytokine refractory population**, an indirect comparison was employed using data from two RCTs. The reliability of an indirect comparison depends on the reliability of the results of the included studies and on the comparability of the trials included.

Regarding the reliability of the results of the included studies, there is some concern about the results for overall survival in the TARGET trial due to treatment switching from placebo to sorafenib at the point of disease progression. As reported in the MS (*Pfizer submission, page 79*):

- *"for OS, the TARGET treatment effect was substantially confounded by switching from the control to treatment arm at the point of progression. While the TARGET trial publication includes a HR which censors those patients who cross over, this approach can lead to severe selection bias if patient's probability of switching treatments is strongly related to their underlying prognosis, which is likely in this setting as patients often switch treatments because their condition has deteriorated." (Source: Pfizer submission, Section 6.7.2, page 79)*

According to the manufacturer this is likely to underestimate the true OS. However, the submission did not provide a more appropriate analysis and used the overall survival benefit of sorafenib versus BSC from the analysis presented in the TARGET study which censored patients at the point of cross-over, which the MS (*Pfizer submission, page 79*) describes as "uncertain and potentially biased".

Comparability of trial populations used in the analyses is not possible, because patient characteristics are not reported for cytokine refractory patients separately in both trials. When the full trial populations are compared the trials are reasonably comparable, the main difference being slightly better MSKCC-scores in the TARGET trial. In addition, the TARGET trial only reported liver and lung metastases, whilst AXIS reported a broader list of metastatic sites.

For the **sunitinib refractory population**, the evidence relies on a simulated treatment comparison, which means that this comparison is not based on randomised treatment allocation, but on a (very sophisticated) comparison of two single treatment arms. Therefore, there is considerable potential for bias in the outcomes of this analysis. A major omission in the STC is that there is no presentation of an associated SE or 95% CI for any of the results. Therefore, we have no idea of the uncertainty of the estimates. The estimates of mean or median PFS or OS and the associated difference between axitinib and placebo should all have been reported with associated variance estimates.

In addition, the reliability of an STC also depends on the reliability of the results of the included studies and on the comparability of the trials included. It assumes that the methods and results of one trial are applicable in the setting of the other.

Regarding the reliability of the results of the included studies, there was also treatment switching in the RECORD-1 trial. However the RECORD-1 trial did include appropriate statistical analysis methods (rank preserving structural failure time models (RPSFT)) to adjust the overall survival estimates for the effect of patients switching treatments. The HR for overall survival used in the submission is the same as the one accepted by NICE in a previous appraisal (everolimus, TA219). This seems a reasonable estimate of the true HR for overall survival in the RECORD-1 trial.

Regarding the comparability of the treatment arms included in the analyses, there are several reasons for concern:

- As reported by the manufacturer, the baseline characteristics for the prior sunitinib patients in the RECORD-1 study were not reported separately; therefore, only characteristics of the whole placebo population could be utilised in the comparison with the AXIS trial patient population *(See: Pfizer submission, Section 6.7.11, page 98).* The treatment arms are reasonably comparable on the patient characteristics reported, the main difference being slightly better ECOG/KPS and MSKCC-scores in the RECORD-1 placebo-arm.

- *"In contrast to AXIS, where all patients included in the study were required to have progressed on first-line therapy by RECIST-defined criteria, in the overall RECORD-1 population, 14% of patients (n=58) discontinued previous TKI therapy because of unacceptable toxicity. Among the subgroup of 58 patients who were intolerant to previous TKI therapy, 45 patients and 13 patients were randomly assigned to everolimus and placebo, respectively. Thus, patients in the RECORD-1 study could have discontinued prior treatment due to intolerance and therefore results would be more reflective of a first-line study." (Source: Pfizer submission, Section 6.7.11, page 108)*
- *"In contrast to the AXIS study, where patients were required to have received only one prior therapy (sunitinib or a cytokine, or bevacuzimab + interferon-α or temsirolimus), patients in the RECORD-1 study were allowed to have received more than one previous therapy and could have been treated with sunitinib or sorafenib, as well as a cytokine in some cases. (Source: Pfizer submission, Section 6.7.2, page 82)*

Regarding this last issue the manufacturer states:

"However, the median OS and patient characteristics have never been reported for those patients in the BSC arm that had progressed on sunitinib after receiving only one line of therapy. The closest available patient populations reporting overall survival data to allow the STC comparison were the ITT BSC population (corrected for crossover using the RPSFT method) and patients receiving everolimus treatment with only prior sunitinib therapy." (Source: Pfizer submission, Section 6.7.11, page 108)

All these issues regarding the clinical effectiveness carry over into the cost effectiveness evaluation.

1.7 Key issues

Overall, the manufacturer's submission is clearly presented and the analyses and underlying assumptions are sound and reasonable. The main issue with this submission is whether a simulated treatment comparison (STC) presents a valid and reliable estimate of the clinical effectiveness of axitinib versus BSC in a sunitinib refractory population. As there is no direct trial evidence it is not possible to compare the results of the STC to any existing evidence so the accuracy and reliability of the results cannot be ascertained. In addition, the uncertainty around the STC results is missing.

The ERG found one error in the indirect comparison (cytokine refractory population) but this did not affect the conclusions. The methods for the STC (sunitinib refractory population) follow the recommended method and it appears to be well-conducted and clearly explained, including considerations of the limitations of the method and available trial evidence. However, the STC is basically a comparison of individual treatment arms and is open to considerable bias. There is no way to assess whether or not the final result is biased.

In addition, the reliability of the results of the included studies and the comparability of the trials included are potential issues. These have been adequately reported in the manufacturer's submissions, and have been summarised in this ERG report.

In conclusion, the ICER for axitinib in a cytokine refractory population is based on a well performed indirect comparison. The results seem reliable, and present an ICER (£64,388 with PAS and [redacted] without PAS) with large confidence intervals due to considerable uncertainty. The ICER for axitinib in a sunitinib refractory population (£40,933 with PAS and [redacted] without PAS) is based on an STC which did not included the uncertainty surrounding the estimates used in the analyses, and is basically a comparison of individual treatment arms and is therefore open to considerable bias.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem.

In Section 2.1 of the manufacturer's submission (*Source: Pfizer Submission, Section 2.1, p23*) Pfizer provided a summary of incidence and prevalence in England and Wales based on credible sources.

> *In the UK, kidney cancer accounts for 3% of male cancers and 2% of female cancers.[6] There are approximately 8163 incident kidney cancers in England and Wales every year[6] of which 7347 are RCC. Of these patients, 27% and 14% are expected to have stage III and IV disease, respectively, and 33% of former stage I-II are expected to recur to stage III-IV, resulting in approximately 4456 patients diagnosed with advanced/mRCC per year (NICE TA169[7, 8] updated with 2009 estimate from the British Association of Urological Surgeons[9]). (Source: Pfizer Submission, Section 2.1, P23).*

Brief evidence was also given of the characteristics of advanced renal cell cancer (aRCC), its aetiology and proposed treatment pathway.

In the manufacturer's submission, RCC is classified according to its histological subtype (e.g. clear cell) and stage.

> *RCC is commonly staged using the American Joint Cancer Committee (AJCC) Tumour Node Metastasis (TNM) staging system. This staging system classifies the size of the tumour (T), the involvement of regional lymph nodes (N) and the presence of distant metastases (M). Advanced RCC, where the tumour is locally advanced or has spread to regional lymph nodes is classed as stage III. Metastatic RCC, where the disease has spread beyond the regional lymph nodes and to distant sites, is classed as stage IV.[8] (Source: Pfizer Submission, Section 2.1, P23).*

Section 2.2 of the manufacturer's submission provides data on the number of patients eligible to receive axitinib each year. Approximately 4,456 patients are diagnosed with advanced/mRCC each year of which 68% are eligible for first-line therapy.

- *Approximately 77% of those eligible for first-line treatment will receive sunitinib (2,333 patients) and approximately 5% cytokines (151 patients) (Pfizer, data on file).*
- *It is estimated that 64% of patients that received sunitinib or cytokines first-line would be eligible for second-line treatment with axitinib (Pfizer, data on file). (Source: Pfizer Submission, Section 2.2, P24).*

Thus, according to the manufacturer's submission, approximately 1,580 patients each year are estimated to be eligible to receive axitinib treatment (1,484 having previously received sunitinib, and 96 having previously received cytokines). It should be noted that these data were not obtained from external sources but derived from Pfizer's 'data on file'.

A description of prognosis and survival is given in Section 2.3 of the manufacturer's submission.

> *The prognosis for advanced/mRCC is poor; the 5-year survival rate is approximately 10%.[10] It is reported that UK patients survived approximately 4 months (median) once they have progressed following treatment with sunitinib[11, 12] and 10 to 13 months for patients who have been treated with cytokines or other agents used prior to the introduction of targeted therapy.[13, 14] Life*

expectancy for these patients is expected to be substantially lower than the 24 months used by NICE to define end of life treatments.(Source: Pfizer Submission, Section 2.3, P24).

Overall, the evidence presented in this section of the submission is in line with the background information given in the final scope.[1] This is also consistent with the ERG's understanding of the problem.

2.2 Critique of manufacturer's overview of current service provision

Section 2.5 of the submission (*Source: Pfizer Submission, Section 2.5, P25*) states that:

There is no cure for advanced/mRCC, therefore the goals of medical intervention are to extend life, prevent worsening of disease, relieve symptoms and maintain physical function.[10] Advanced/mRCC is largely resistant to chemotherapy, radiotherapy and hormonal therapy.[10]

Advances in the understanding of the molecular biology of RCC have led to the development of targeted therapies. Current targeted agents have focused on two pathways that are commonly de-regulated in RCC, the VEGFR pathway (e.g. sunitinib and pazopanib) and the mammalian target of rapamycin (mTOR) pathway (e.g. temsirolimus and everolimus).

Current first-line treatment options in the UK include the TKIs, sunitinib and pazopanib, both of which have received a positive recommendation from NICE.[8, 15] NICE does not currently recommend any interventional therapies for advanced/mRCC following failure of initial systemic therapy and patients subsequently receive BSC (defined as the provision of drug and non-drug therapy for the relief of symptoms and general patient management.[16] (Source: Pfizer Submission, Section 2.5, P25).

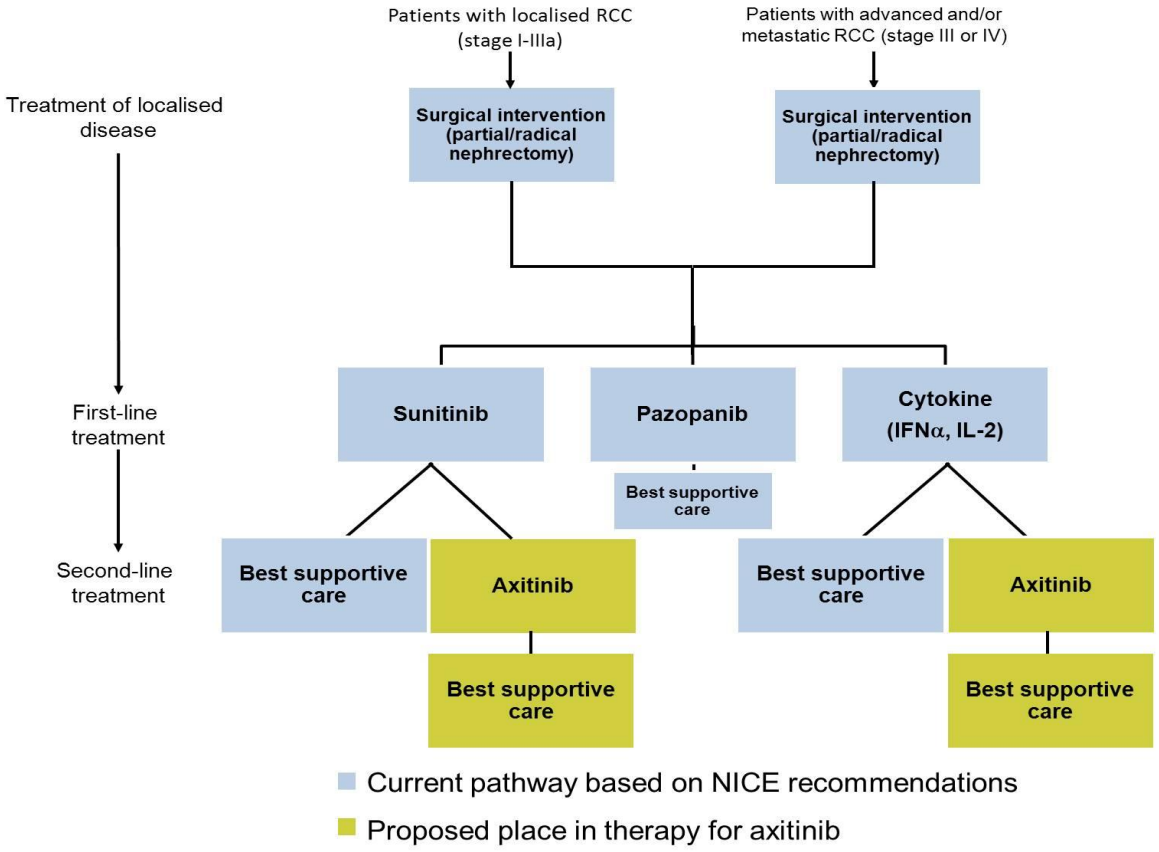
The manufacturer's submission notes that in 2009 NICE issued guidance (TA169) recommending sunitinib for the first-line treatment of advanced and/or metastatic RCC in patients who were suitable for immunotherapy and with an ECOG performance status of 0 or 1.[8] Bevacizumab, sorafenib and temsirolimus have also undergone a NICE assessment in the form of a multiple technology appraisal (MTA) with the final guidance issued in August 2009 (TA178).[10] The use of bevacizumab, sorafenib or temsirolimus was not recommended for first line. In addition, use of sorafenib or sunitinib was not recommended for second line.[10] In February 2011, NICE issued guidance (TA215) recommending pazopanib as a first-line treatment option for patients with advanced RCC who had not received prior cytokine therapy and with an ECOG performance status of 0 or 1.[15] In April 2011, everolimus received a negative recommendation for the second-line treatment of patients with advanced RCC (TA219).[16]

In summary:

NICE recommends sunitinib or pazopanib for the first-line treatment of patients with advanced/mRCC with an ECOG performance status of 0 or 1. NICE does not currently recommend any interventional therapies for advanced/mRCC following failure of initial systemic therapy (Source: Pfizer Submission, Section 2.4, P25).

The manufacturer proposes a treatment pathway for patients with advanced/mRCC in England and Wales, which is shown in the figure below (Figure 2.1); this is based on NICE guidance issued to date and the place in therapy of axitinib as anticipated by Pfizer. *(Source: Pfizer Submission, Section 2.5, P26).*

Figure 2.1: Proposed treatment pathway based on current NICE recommendations for patients with advanced/metastatic renal cell carcinoma *(Source: Pfizer Submission, Section 2.5, P26)*



There are no therapies currently recommended by NICE for second-line treatment of patients with advanced/mRCC for whom first-line therapy has failed.[10] Therefore the main comparator in this submission is best supportive care (BSC) in line with the scope and current NICE guidance. (*See: Pfizer Submission, Section 2.7, P27*)

The manufacturer notes that axitinib can be taken at home and therefore its administration will not incur any additional resource use. However, patients may need to be treated for adverse events associated with axitinib. The most common adverse events experienced by patients treated by axitinib are diarrhoea (54.9% of patients), hypertension (40.4% of patients) and fatigue (39% of patients). (*See: Pfizer Submission, Section 2.8 and 2.9, P27*)

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

3.1 Population

The population considered by the submission is:

> *Adult patients with advanced renal cell carcinoma (aRCC) after failure of prior treatment with sunitinib or a cytokine. (Source: Pfizer Submission, Section 5, P31).*

This is an adequate description of the population under consideration, although it differs slightly with that defined in the NICE scope (prior systemic treatment). However, this was done to meet the current licensing requirements. Overall, the ERG agrees that the population considered is reflective of the actual clinical population.

3.2 Intervention

The intervention is axitinib (Inlyta®, Pfizer) the recommended starting dose is 5mg twice daily. Section 1.2 in the submission (*Pfizer Submission, Section 1.2, P16*) states:

> *Axitinib is the first, next-generation, oral vascular endothelial growth factor receptor tyrosine kinase inhibitor (VEGFR-TKI). Axitinib selectively inhibits the VEGF receptors (VEGFR)-1, -2, and -3 with greater potency and selectivity than currently available VEGFR-TKIs (Source: Pfizer Submission, Section 1.2, P16).*

Axitinib is not commercially available in the UK but according to the manufacturer's submission the anticipated date for commercial availability in UK is October 2012. The intervention is in line with the final scope.

3.3 Comparators

The single comparator mentioned in the NICE scope was best supportive care (BSC). In the manufacturer's submission placebo was used as a proxy for BSC (*Source: Pfizer Submission, Section 6, P34*)

BSC seems an appropriate comparator as there is no access to another funded second-line treatment for the UK patient population and it is the only comparator specified in the NICE scope. Furthermore, Pfizer's choice to use placebo as a proxy for BSC seems reasonable. In the PenTAG MTA report for the assessment of bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma, the authors concluded:

> *We were unable to find any useful definitions of BSC in this population in the literature, or any trials that compare sorafenib or sunitinib with BSC. ... We have therefore assumed that treatment with placebo is equivalent to BSC. (Source: PenTAG Assessment Report for TA179, P41[4])*

3.4 Outcomes

The outcome measures mentioned in the NICE scope are overall survival, progression free survival, response rates, adverse effects of treatment and health-related quality of life. These outcomes are all reported in the AXIS trial comparing axitinib with sorafenib. However, for the comparison axitinib with BSC, only overall survival and progression free survival were reported in the manufacturer's submission.

The outcomes for the economic analysis were incremental cost per quality-adjusted life year. Costs were considered from the NHS and Personal Social Service perspective.

3.5 Other relevant factors

The section on equality issues has been marked 'not applicable' by the manufacturer. (*Source: Pfizer Submission, Section 3, P28*).

This report is based on the manufacturer submission that was received on 24 September 2012 and is the final version with PAS and updated CiC. Page numbers mentioned in this report should be based on this version of the MS. The ERG received five versions of the MS in total. Therefore, there may be some confusion when referencing page numbers from the MS. However, section numbers were the same in all versions of the MS as far as we know.

Sub group analysis: the final scope states that if the there is enough evidence two sub group analyses should be considered namely for prior treatment and for prognostic score (for example, ECOG or Motzer).[1] In the submission a subgroup analysis was done for adult patients with advanced/mRCC after failure of prior treatment with sunitinib or a cytokine. No subgroup analyses were done for prognostic score. This was justified as follows:

Whilst PFS for the total population of patients included in the AXIS trial has been sub-analysed by performance status, this analysis has not been conducted for the sub-population of patients after failure of prior treatment with sunitinib or a cytokine because the resulting sub-groups are too small for interpretable results (Source: Pfizer Submission, Section 5, P31).

This seems reasonable.

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods used by the manufacturer to systematically review clinical effectiveness evidence

4.1.1 State objective of systematic review. Provide description of manufacturer's search strategy and comment on whether the search strategy was appropriate. If the manufacturer did not perform a systematic review, was this appropriate?

An evidence based checklist for the Peer Review of Electronic Search Strategies (PRESS), developed by McGowan et al. was used to inform this critique.[17] The submission was checked against the Single Technology Appraisal (STA) specification for manufacturer/sponsor submission of evidence.[18] The ERG has presented only the major limitations of each search strategy in the main report. Further criticisms of each search strategy can be found in Appendix 1B.

1. Clinical effectiveness

Searches were carried out on all databases required by NICE (MS 6.1 & 10.2[2]). The search dates for the original and update searches were reported for the majority of resources. The MS failed to report the host, indices used (i.e. Science Citation Index), the date range searched, and search date for the Web of Science search. The ERG queried these omissions in the Points of Clarification letter and these were addressed in the manufacturer's response.[19] The ERG noted a few general areas of weakness in the strategies for this section (Appendix 1B); however, it is unlikely that these errors would have impacted in the recall of results.

The manufacturer reported that the following additional searches were undertaken for this section; the FDA website was searched for reports published by the Oncologic Drugs Advisory Committee. Hand searches were undertaken for the conference proceedings of the American Society of Clinical Oncology: ASCO (including the genitor-urinary symposium ASCO-GU), the European Society for Medical Oncology (ESMO) and the European Cancer Organisation (ECCO). The MS did not include details of the search terms used to search these additional resources; therefore the ERG was unable to comment on these searches due to insufficient information. Finally clinical study reports were provided by the manufacturer.

In section 6.1 the MS also reported an additional set of searches intended to "*identify RCTs and Non-RCTs reporting efficiency and safety data for patients with mRCC who received BSC following progression with first-line sunitinib treatment*" (MS 10.2[2]). These searches were conducted on Medline, Medline in Process, Embase and Cochrane Library. The ERG noted that the strategies in this section contained some redundant lines; however, these errors were not thought to be consequential; for further details please see Appendix 1B. Additional resources searched in this section included the ASCO, ASCO-GU, ESMO and ECCO conferences, as well as a range of trials and specialist cancer trials resources including ClinicalTrials.gov.

In conclusion the manufacturer translated the research question into appropriate search strategies and the ERG considered these searches to be adequate.

Indirect and mixed treatment comparisons

The MS reported that the strategies reported in 10.2 were employed for this section. Therefore the same limitations applied to these searches.

Adverse events

The Manufacturer stated that searches created for sections 6.1 were also designed to identify eligible studies for adverse events associated with axitinib. CRD guidance recommends that if searches have been limited by an RCT filter, additional searches should be undertaken to ensure that adverse events that are long-term, rare or unanticipated are not missed.[20] Despite the addition of a Systematic Reviews and HRQL filter the ERG considered that it was possible that some relevant evidence might not have been identified as a consequence of the study design limits. However this may have been mitigated by searches of the additional resources listed in 10.2.

Non-RCT Evidence (Axitinib)

Adequate searches were carried out on all NICE required databases (MS 10.6[2]). The MS also reported the use of additional information derived from clinical study report provided by the manufacturer. The ERG noted the same limitations in the facet for axitinib as in earlier searches (see Appendix 1B).

2. Cost effectiveness

Searches were carried out on all NICE required databases (MS 10.10[2]). The search dates were reported for all searches. No host was reported for the Embase or EconLit searches, but the ERG presumed this to be Ovid, as the search syntax matched the Ovid Medline strategy. The searches were well reported and reproducible. There was some disparity between the way that renal cell carcinoma was searched for in this and the HRQL search, compared with the earlier clinical effectiveness searches, but it is not clear if this would have impacted the overall recall of the strategies. The ERG was concerned that the NHS EED search contained an economics facet. NHS EED is a database which is specifically made up of economic evaluations and therefore this filter was not necessary. However following an investigation of this search using the Wiley interface, the ERG concluded that this would have had little or no impact on the results. The ERG noted that searches in this section were limited to English language only, which may have resulted in the omission of potentially useful papers in other languages.

Additional searches were also reported for ASCO, ESMO, ECCO and the International Society for Pharmacoeconomics and outcomes research (ISPOR) conferences for 2011-2012. The MS did not include details of the search terms used to search these additional resources; therefore, the ERG was unable to comment on these searches. The MS also reported a search of the Cost Effectiveness Analysis Registry using keywords for "renal cell carcinoma" and "RCC" and a search of the NICE website for relevant documents for second-line RCC treatments. Further hand searching of references lists of included publications and other relevant reviews was also reported.

Measurement and valuation of health effects

Searches were carried out on all NICE required databases (MS 10.12[2]). Searches were adequate and easily reproducible. The ERG noted the same error in the application of a HRQL filter to the NHS EED search as with the cost effectiveness search. After an investigation of this search using the Wiley interface, the ERG concluded that this would have had little or no impact on the results. Search strategies in this section also appeared to be missing the HRQL outcome SF-6D. Any papers missed by this omission may have been retrieved by the HRQL facet in the earlier Clinical Effectiveness searches (10.2), which did include this outcome. The ERG noted that searches in this section were limited to English language only, which may have resulted in the omission of potentially useful papers in other languages. Again this may have mitigated by the earlier results retrieved by the Clinical Effectiveness searches.

Additional searches were reported of the following conference abstracts from 2011-2012: ASCO, ASCO-GU, ESMO and ECCO. The following resources were also searches for the keywords "renal cell carcinoma" and "RCC"; Cost Effectiveness Analysis Register, EQ-5D website and Research papers in Economics (RePEc)

Resource identification, measurement and valuation

The ERG noted that in the NICE specification section 7.5.3, NICE requests a systematic search of relevant source data for resource identification. In the MS submission it was stated that a systematic review was not completed and that evidence for this section came from clinical opinion, published sources, the manufacturer's submission for the everolimus STA and the PenTAG model developed for the NICE bevacizumab, sorafenib, sunitinib and temsirolimus MTA. The ERG queried this omission in the Points of Clarification letter. The manufacturer responded:

> "*As multiple recent appraisals have been carried out recently in mRCC, with several opportunities for consensus and input from NICE ERGs and appraisal committees, it was felt than an updated systematic review would not be required. Furthermore, assuming common resource utilisation assumptions with previous NICE mRCC appraisals ensures consistency in decision-making. To ensure the validity of the previous NICE assumptions with current UK practice, resource utilisation assumptions referenced from previous appraisals were tested with UK clinical experts in mRCC treatment to ensure that they were still relevant and reflective of clinical practice prior to their being incorporated into the submission".*[19]

Summary of searching

The searches documented in the initial manufacturer's submission contained several areas of weakness, two of which relating to reproducibility were queried by the ERG. The manufacturer addressed all the points of concern raised by the ERG in their response to clarification. Despite these weaknesses the ERG concluded that searching was carried out to an adequate standard and accurately reflected the research questions.

4.1.2 State the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

The inclusion/exclusion criteria used in the study selection are described in Table 4.1, page 37 of the MS (*see Pfizer's submission, Section 6.2.1, page 37-38*; and the Table below).

Table 4.1: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Description	Justification
Inclusion criteria		
Population	Adult patients with metastatic RCC who have received first- or second-line treatment.	Patients had received prior systemic therapy, as specified in the NICE scope.
Interventions	Any chemotherapy or targeted therapy in the second-line setting (RCT search only) Axitinib in the second-line setting (non-RCT search only)	In addition to the comparator stated in the scope (BSC), other interventions (both first and second-line) were searched in the systematic review. Studies where patients received a therapy as first-line treatment were later excluded for the purpose of this submission.
Outcomes	**Efficacy:** OS, PFS, TTP, ORR (complete + partial response), Proportion of patients with stable disease, Duration of response, Time to response, Symptom assessments (where reported), Time to deterioration (composite/individual endpoint) **Safety:** Incidence and severity of AEs. Quality of life or any other global patient-reported outcomes	Consistent with final scope
Study design	Prospective randomised controlled trials (for the RCT search) Non-RCTs (for the non-RCT search)	Separate searches were conducted for RCTs and non-RCTs
Language restrictions	English language only	To reduce number of hits and to identify studies in patient populations relevant to the UK setting
Exclusion criteria		
Population	Subjects <18 years of age	As specified by final scope
Interventions	Radiotherapy, surgery and other non-relevant comparators	Not relevant to final scope
Outcomes	Studies not investigating efficacy, safety or QoL	Not relevant to final scope
Study design	Non-RCTs (for the RCT search) RCTs (for the non-RCT search)	Separate searches were conducted for RCTs and non-RCTs
Language restrictions	Abstracts published in non-English language	To reduce number of hits and to identify studies in patient populations relevant to the UK setting

Abbreviations: AE: adverse event; BSC, best supportive care; NICE, National Institute for Health and Clinical Excellence; ORR, objective response rate; OS: overall survival; PFS: progression-free survival; QoL: quality of life; RCC: renal cell carcinoma; RCT: randomised controlled trial; TTP: time to progression

Comment

The inclusion criteria are consistent with the NICE scope. The only questionable element is the use of language restrictions. However, the ERG is not aware of any foreign language publications meeting the inclusion criteria.

The manufacturer performed additional searches to identify clinical studies (RCTs and non-RCTs) which reported efficacy and safety data in patients with mRCC who received BSC following

progression with first-line sunitinib treatment (*see Pfizer's submission, Section 10.15, page 318-327*). The inclusion and exclusion criteria were described in Table 4.2.

Table 4.2: Inclusion and exclusion criteria for additional searches

Inclusion criteria	
Population	Adult patients with metastatic RCC
Age	≥ 18 years
Gender	No restrictions
Race	No restrictions
Outcomes	Including but not restricted to: Overall survival (OS) Progression-free survival (PFS) Time to progression (TTP) Survival after progression (SAP) Survival measures that reported on 1st and 2nd line treatment were collected
Interventions	First-line therapy: Sunitinib Cytokine therapy (IL-2, IFNa) Second-line therapy: Best supportive care†
Comparators	No restrictions
Study design	Prospective, randomised controlled trials Prospective non-randomised controlled studies
Language	No restrictions

†There is currently no standard definition of what treatments constitute best supportive care. A commonly used definition is 'any palliative therapeutic modality that may be offered to the patient excluding chemotherapy but including radiotherapy and non-cytotoxic medication'. This includes antibiotics, analgesics, antiemetics, corticosteroids, blood transfusions, nutritional support and focal external-beam radiation for control of pain, cough, dyspnoea or haemoptysis.

Comment

The inclusion criteria are again consistent with the NICE scope. However, the purpose of this search was to construct a network linking axitinib with placebo in the sunitinib refractory population. Therefore, the inclusion of cytokine therapy as first-line therapy seems odd and, more importantly, restricting second-line therapy to 'best supportive care' is a serious mistake as this would exclude studies comparing sorafenib with placebo in a sunitinib refractory population.

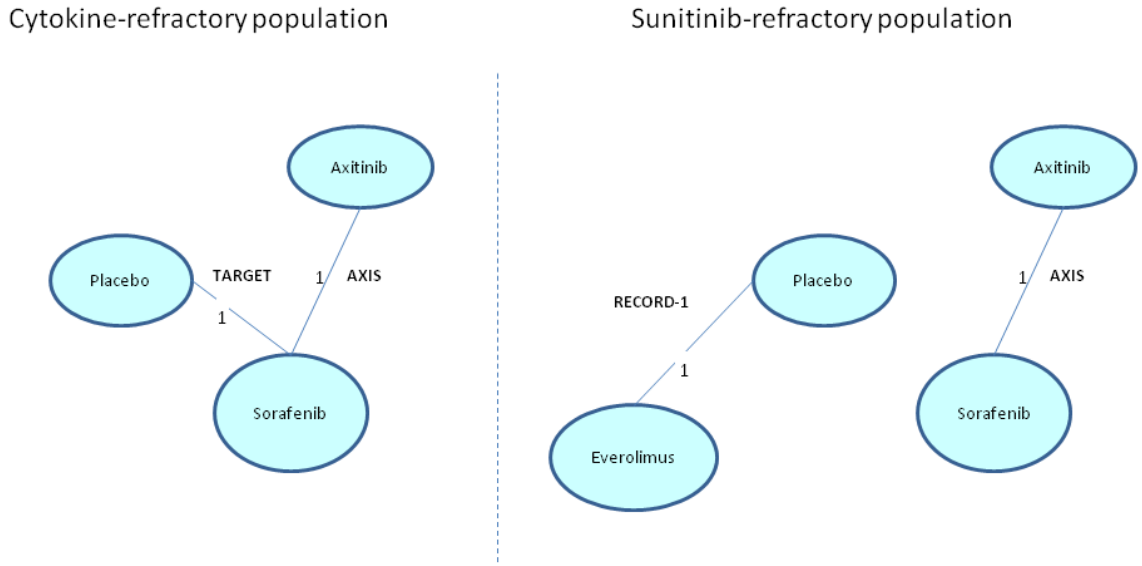
The ERG performed their own searches to try to bridge the gap between axitinib and BSC/placebo in the sunitinib refractory population. We found one study comparing temsirolimus with sorafenib in a sunitinib refractory population.[21] As this study is sponsored by Pfizer and the primary completion date is January 2012 Pfizer should be able to access data from this study. This does add another treatment that can be compared indirectly with axitinib. However, as we were not able to link temsirolimus with either everolimus or placebo in the relevant population, we were also not able to bridge the gap.

4.1.3 What studies were included in the clinical effectiveness review and what were excluded? Provide a table of identified studies. Please identify the <u>most important</u> clinical effectiveness studies.

For the comparison between axitinib and BSC in a cytokine refractory population, two trials were included: AXIS and TARGET. Similarly, for the comparison between axitinib and BSC in a sunitinib

refractory population, two trials were included: AXIS and RECORD-1. The evidence network is described in Figure 19 of the MS (*MS, Section 6.7.2, page 78*) – see figure below.

Figure 4.1: Evidence networks for the cytokine-refractory and sunitinib refractory populations
(Source: Pfizer Submission, section 6.7.2, page 78)



In total three trials were included, using data from two specific populations: cytokine refractory and sunitinib refractory patients:

- AXIS: a Phase III study of axitinib vs sorafenib[22]
- TARGET: a Phase III study of sorafenib vs placebo[23, 24]
- RECORD-1: a Phase III study of everolimus vs placebo[25, 26]

The study and patient characteristics are presented in Table 4.3. Patient characteristics were reported for all three full trials. However, patient characteristics for the subgroups of cytokine refractory and sunitinib refractory patients are not reported in most cases. Patient characteristics for cytokine refractory patients are completely missing for the AXIS and TARGET trials; while patient characteristics for sunitinib refractory patients are missing for the sorafenib arm in the AXIS trial and for the placebo arm in the RECORD-1 trial.

Table 4.3: Comparison of the study and patient characteristics in the three included trials: AXIS, TARGET and RECORD-1

	AXIS	TARGET	RECORD-1
Trial methodology			
Intervention and comparator	Axitinib (N=361) 5 mg BD starting dose Sorafenib (N=362) 400 mg BD starting dose	Sorafenib (N=451) 400 mg BD Placebo (N=452) BD	Everolimus (N=272) 10 mg OD starting dose Placebo (N=138)
Population	Patients with mRCC following failure of a prior systemic first-line regimen containing one of the following: sunitinib, bevacizumab + IFNα, temsirolimus or cytokine(s).	Patients with metastatic clear cell RCC who had progressed after one prior systemic therapy in the previous 8 months	Patients with metastatic clear cell RCC who had progressed after sunitinib or sorafenib, or both. Previous therapy with bevacizumab, interleukin 2, or interferon alfa was also permitted.
Inclusion criteria	• ≥ 18 years • Histologically/ cytologically confirmed mRCC with a clear cell subtype component • Evidence of measurable disease (by RECIST) • Progressive disease criteria per RECIST (Version 1.0) after 1 prior systemic first-line regimen for mRCC. The prior regimen had to have contained 1 of the following: sunitinib, bevacizumab + IFN-α, temsirolimus, or cytokine(s) • ECOG performance status of 0 or 1 • Life expectancy of ≥ 12 weeks • At least 2 weeks since the end of prior systemic treatment (4 weeks for bevacizumab + IFNα) • Adequate renal, hepatic and haematological function	• ≥ 18 years • Histologically confirmed metastatic clear cell RCC which had progressed after 1 systemic treatment • ECOG PS ≤ 1 • MSKCC favourable or intermediate risk • Life expectancy of ≥ 12 weeks • Adequate bone marrow, liver, pancreatic and renal function • Prothrombin time of or partial thromboplastin time <1.5 x ULN	• ≥ 18 years • Evidence of measurable disease (by RECIST) • KPS ≥ 70% • Adequate bone marrow, hepatic and renal function
Exclusion criteria	• Prior treatment of mRCC with more than 1 systemic first-line regimen • History of malignancy other than RCC • A need for CYP3A4 inhibiting/inducing or CYP1A2 inducing drugs • CNS metastases • Uncontrolled hypertension	• Brain metastases • Previous exposure to VEGF inhibitors	• Previous mTOR inhibitor therapy (temsirolimus), • Untreated CNS metastases, • Uncontrolled medical conditions (eg, unstable angina pectoris, symptomatic congestive heart failure, recent myocardial infarction, or diabetes).

	• Myocardial infarction, uncontrolled angina, congestive heart failure or cerebrovascular accident in previous 12 months • DVT or pulmonary embolism in previous 6 months		
Design	Randomised, multicentre, international Phase III study. Cross-over was not permitted.	Randomised, double-blind, placebo-controlled, Phase III study. Cross-over was allowed following the first PFS analysis.	Randomised, double-blind, placebo-controlled, Phase III study. Cross-over was allowed following disease progression.
Duration of study	Treatment was to continue until disease progression, intolerable adverse drug reactions or withdrawal of consent.	Until disease progression or withdrawal due to AEs.	Until disease progression, unacceptable toxicity, death, or discontinuation for any other reason.
Method of blinding	Open-label, however the independent assessment of the primary endpoint (PFS) was done in a blinded manner by the IRC	Double-blind	Double-blind
Tumour assessments	CT/MRI and bone scans were performed at screening, at 6 weeks and 12 weeks, then every 8 weeks thereafter.	Progression of disease was determined by CT or MRI, clinical progression or death by RECIST. Assessments of responses required confirmatory findings on CT or MRI 4 or more weeks after the initial determination of a response.	Tumour measurements (assessed by CT or MRI scans) were done at screening and were subsequently repeated every 8 weeks for the remainder of the study, as well as on discontinuation of study drug.
Duration of follow-up	Patients were followed until disease progression, intolerable adverse drug reactions or withdrawal of consent.	Until disease progression or withdrawal due to AEs, until death.	Until disease progression, unacceptable toxicity, death, or discontinuation for any other reason.

Patient characteristics

	AXIS		**TARGET**		**RECORD-1**	
	Axitinib	**Sorafenib**	**Sorafenib**	**Placebo**	**Everolimus**	**Placebo**
All Patients	**N=361**	**N=362**	**N=451**	**N=452**	**N=272**	**N=138**
Age, median (range)	61 (20-82)	61 (22-80)	58 (19-86)	59 (29-84)	61 (27-85)	60 (29-79)
Sex, % male	73%	71%	70%	75%	78%	76%
ECOG/KPS						
- 0 / 90-100	54%	55%	49%	46%	64%	67%
- 1 / 70-80	45%	44%	49%	52%	36%	33%
- 2 / 50-60	<1%	0	2%	1%	0	0
- missing	0	0	<1%	<1%	<1%	0

MSKCC - Favourable (0) - Intermediate(1) - Poor (≥1) - missing/NA	28% 37% 33% 2%	28% 36% 33% 3%	prognostic risk Low: 52% Intermediate: 48% 0	50% 49% <1%	29% 56% 15%	28% 57% 15%
Previous nephrectomy	91%	91%	94%	93%	96%	95%
Previous radiotherapy	NR	NR	27%	24%	31%	28%
Clear cell RCC	98%	99%	100%	100%	100%	100%
Metastatic sites - Lung - Liver - Bone - Lymph node	 76% 28% 33% 58%	 81% 29% 30% 56%	 77% 26%	 77% 26%	 73% 35% 37% 75%	 81% 36% 31% 71%
Previous cytokine	126 (35%)	125 (35%)	374 (83%)	368 (81%)	unknown (>0)	unknown (>0)
Previous sunitinib	194 (54%)	195 (54%)	0	0	127 (47%)	unknown (>0)
	AXIS		**TARGET**		**RECORD-1**	
	Axitinib	**Sorafenib**	**Sorafenib**	**Placebo**	**Everolimus**	**Placebo**
First-line cytokine	126 (35%)	125 (35%)	374 (83%)	368 (81%)	--	--
Age, median (range)	NR	NR	NR	NR		
Sex, % male	NR	NR	NR	NR		
ECOG/KPS - 0 / 90-100 - 1 / 70-80 - 2 / 50-60 - missing	NR	NR	NR	NR		
MSKCC - Favourable (0) - Intermediate(1) - Poor (≥1)	NR	NR	NR	NR		

- missing/NA						
Previous nephrectomy	NR	NR	NR	NR		
Previous radiotherapy	NR	NR	NR	NR		
Clear cell RCC	NR	NR	NR	NR		
Metastatic sites - Lung - Liver - Bone - Lymph node	NR	NR	NR	NR		
Weeks on cytokine	NR	NR	NR	NR		
Previous sunitinib	NR	NR	NR	NR		
	AXIS		**TARGET**		**RECORD-1**	
	Axitinib	**Sorafenib**	**Sorafenib**	**Placebo**	**Everolimus**	**Placebo**
First-line sunitinib	194 (54%)	195 (54%)	--	--	127 (47%)	139*
Age, median (range)	61 (22-82)	NR			59 (28-81)	60 (29-79)
Sex, % male	74%	NR			80%	76%
ECOG/KPS - 0 / 90-100 - 1 / 70-80 - 2 / 50-60 - missing	 52% 48% 0 0	NR			 60% 41% 0 1%	 68% 33% 0 0
MSKCC - Favourable (0) - Intermediate(1) - Poor (≥1) - missing/NA	 20% 41% 36%	NR			 28% 55% 17%	 28% 57% 15%
Previous nephrectomy	88%	NR			91%	95%
Previous radiotherapy	23%	NR			31%	27%

Clear cell RCC	98%	NR			100%	100%
Metastatic sites - Lung - Liver - Bone - Lymph node	NR	NR			NR	 <u>81%</u> <u>38%</u> <u>30%</u> <u>70%</u>
Weeks on sunitinib	41.4 (2.7-471)	NR			41.3 (1.3-120)	NA
Previous cytokine	0	NR			unknown (>0)	unknown (>0)

ECOG=Easter Cooperative Oncology Group, KPS=Karnofsky performance status, MSKCC=Memorial Sloan-Kettering Cancer Centre, RCC=renal cell carcinoma, NA=Not applicable, NR=Not reported.

*) These data are from the Pfizer submission (*MS, Table 21, page 101*) and for the placebo arm include all patients having received sunitinib or sorafenib previously, some received up to 4 previous treatments. According to the Pfizer submission, 21% (n=89) of patients in RECORD-1, had received only one prior systematic therapy, sunitinib or sorafenib. The number of patients in the placebo arm having received only prior sunitinib is not known. According to Motzer et al. 2008[26], 124 out of 272 (46%) in the everolimus arm and 60 out of 138 (43%) in the placebo arm had received sunitinib only as prior systemic treatment. Some of the data are only reported in Motzer et al. 2010[25], these have been added for the placebo group <u>underlined</u>.

4.1.4 Provide details of any relevant studies not discussed in the submission? Why were these studies excluded and how were these studies identified by the ERG?

The ERG is not aware of any relevant studies that were not included in the MS.

4.2 Summary and critique of submitted clinical effectiveness evidence

4.2.1 Summary of submitted clinical evidence for each relevant trial.

In total three trials were included, using data from two specific populations: cytokine refractory and sunitinib refractory patients:

- AXIS: a Phase III study of axitinib vs sorafenib[22]
- TARGET: a Phase III study of sorafenib vs placebo[23, 24]
- RECORD-1: a Phase III study of everolimus vs placebo[25, 26]

The study and patient characteristics are presented in section 4.1.3. The results are presented in the Table below.

Table 4.4: Results of the three included trials: AXIS, TARGET and RECORD-1

	AXIS		TARGET		RECORD-1	
	Axitinib	**Sorafenib**	**Sorafenib**	**Placebo**	**Everolimus**	**Placebo**
All Patients	**N=361**	**N=362**	**N=451**	**N=452**	**N=272**	**N=139**
Overall survival	(95% CI)					
- Death	211 (58%)	214 (59%)	171 (38%)	196 (43%)[j]	NR	NR
- Median time to	NR	NR	19.3 m	15.9 m[j]	64.1 wks (14.8 m)	53.4 wks (14.4 m)[f]
- Hazard ratio	axitinib vs sorafenib:	0.969 (0.800, 1.174)[a]	sorafenib vs placebo:	0.77 (0.63, 0.95)[j]	everolimus vs placebo	0.87 (0.65, 1.15)[f]
Progression free	survival (95% CI)					
- PFS event	192 (53%)	210 (58%)	NR	NR	193 (71%)	109 (79%)[g]
- Median time to	6.7m (6.3, 8.6)	4.7m (4.6, 5.6)	5.5 m	2.8 m[j]	4.9 m (4.0, 5.5)	1.9 m (1.8, 1.9)
- Hazard ratio	axitinib vs sorafenib:	0.665 (0.544, 0.812)[a]	sorafenib vs placebo:	0.51 (0.43, 0.60)[j,k]	everolimus vs placebo	0.33 (0.25, 0.43)
Response rates						
- OR event	19.4% (15.4, 23.9)	9.4% (6.6, 12.9)	10% (7, 13)	2% (1, 4)[l]	1.8%	0%
- Duration	11 m (7.4, ne)	10.6 m (8.8, 11.5)	182 days (36, 378)	NR	NR	NR
- OR risk ratio	axitinib vs sorafenib:	2.056 (1.408, 3.003)	NR	NR	NR	NR
Health-related	quality of life					
- TTD[b]	Hazard ratio (95%	0.829 (0.701, 0.981)	NR	NR	NR	NR
- FKSI-DRS	CI):	0.12 (-0.45, 0.69)	NR	NR	Post-Tx difference:	0.82 (0.57, 1.18)
- EQ-5D	Post-Tx difference:	NR	NR	NR	NR	NR
	NR					
Adverse events	All grades[m] Grade	All grades Grade 3/4	All grades Grade	All grades Grade	All grades Grade	All grades Grade
- stomatitis	3/4	20% 1%	3/4	3/4	3/4	3/4
- rash	20% 1%	32% 4%	NR	NR	44% 4%	8% 0%
- fatigue	13% <1%	32% 5%	40% 1%[h]	16% <1%[h]	29% 1%	7% 0%
- asthenia	39% 11%	14% 3%	37% 5%	28% 4%	31% 5%	27% 3%
- diarrhoea	21% 5%	53% 7%	NR	NR	33% 3%	23% 4%
- hypertension	55% 11%	29% 11%	43% 2%	13% 1%	30% 1%	7% 0%

- decr. appetite	40% 16%	29% 4%	17% 4%	2% <1%	NR	NR
- nausea	34% 5%	22% 1%	NR	NR	NR	NR
- dysphonia	32% 3%	14% 0%	23% <1%	19% 1%	26% 1%	19% 0%
- hypothyroidism	31% 0%	8% 0%	NR	NR	NR	NR
- PPE	19% <1%	51% 16%	NR	NR	NR	NR
- alopecia	27% 5%	32% 0%	30% 6%[i]	7% 0%[i]	NR	NR
- infections	4% 0%	NR	27% <1%	3% 0%	NR	NR
- non-inf pneum.	NR	NR	NR	NR	37% 10%	18% 1%
	NR		NR	NR	14% 4%	0% 0%
First-line cytokine	126 (35%)	125 (35%)	374 (83%)	368 (81%)	--	--
Overall survival	(95% CI)					
- Death	51 (41%)	57 (46%)	NR	NR		
- Median time to	NR	NR	NR	NR		
- Hazard ratio	axitinib vs sorafenib:	0.813 (0.555, 1.191)	sorafenib vs placebo:	0.78 (0.62, 0.97)[d]		
Progression free	survival (95% CI)					
- PFS event	50 (40%)	69 (55%)	NR	NR		
- Median time to	12.1m (10.1, 13.9)	6.5m (6.3, 8.3)	NR	NR		
- Hazard ratio	axitinib vs sorafenib:	0.464 (0.318, 0.676)	sorafenib vs placebo:	0.44 (0.35, 0.55)[c]		
Response (95%	CI)					
- OR event	32.5% (24.5, 41.5)	13.6% (8.1, 20.9)	NR	NR		
- Duration	11m (7.4, ne)	10.6 (5.9, 11.5)	NR	NR		
- OR risk ratio	axitinib vs sorafenib:	2.392 (1.434, 3.992)	NR	NR		
Health-related	quality of life					
- TTD[b]	NR	NR	NR	NR		
- FKSI-DRS[n]	Post-Tx difference:	[redacted]	NR	NR		
- EQ-5D[n]	Post-Tx difference:	[redacted]	NR	NR		
Adverse events	TEAEs[o]	TEAEs[o]	NR	NR		
- stomatitis	[redacted]	[redacted]				
- rash	[redacted]	[redacted]				

- fatigue - asthenia - diarrhoea - hypertension - decr. appetite - nausea - dysphonia - hypothyroidism - PPE - alopecia - infections - non-inf pneum.	[redacted] NR NR	[redacted] NR NR				
First-line sunitinib	194 (54%)	195 (54%)	--	--	127 (47%)	139*
Overall survival - Death - Median time to - Hazard ratio	(95% CI) 131 (68%) 65.9 wks (15.2 m) axitinib vs sorafenib:	 131 (67%) NR 0.997 (0.782, 1.270)			 NR 54.4 wks (12.6 m) NR	 NR 43.4 wks (10.0 m) NR
Progression free - PFS event - Median time to - Hazard ratio	survival (95% CI) 117 (60%) 4.8m (4.5, 6.4) axitinib vs sorafenib:	 120 (62%) 3.4m (2.8, 4.7) 0.741 (0.573, 0.958)			 NR 16.9 wks (3.9 m) everolimus vs placebo	 NR 7.8 wks (1.8 m) 0.34 (0.23, 0.51)[e]
Response rates - OR event - Duration - OR risk ratio	 11.3% (7.2, 16.7) 11m (5.2, ne) axitinib vs sorafenib:	 7.7% (4.4, 12.4) 11.1 (ne, ne) 1.477 (0.792, 2.754)			 NR NR NR	 NR NR NR
Health-related - TTD[b] - FKSI-DRS[n,o]	quality of life NR Post-Tx difference:	 NR [redacted]			 NR NR	 NR NR

- EQ-5D[n,o]	Post-Tx difference:	[redacted]			NR	NR
Adverse events	TEAEs[o]	TEAEs[o]			NR	NR
- stomatitis	[redacted]	[redacted]				
- rash	[redacted]	[redacted]				
- fatigue	[redacted]	[redacted]				
- asthenia	[redacted]	[redacted]				
- diarrhoea	[redacted]	[redacted]				
- hypertension	[redacted]	[redacted]				
- decr. appetite	[redacted]	[redacted]				
- nausea	[redacted]	[redacted]				
- dysphonia	[redacted]	[redacted]				
- hypothyroidism	[redacted]	[redacted]				
- PPE	[redacted]	[redacted]				
- alopecia	[redacted]	[redacted]				
- infections	NR	NR				
- non-inf pneum.	NR	NR				

CI=Confidence Interval, ECOG PS=Eastern Cooperative Oncology Group Performance Status, FKSI-15=The 15-item Functional Assessment of Cancer Therapy Kidney Symptom Index, FKSI-DRS=specifically measures symptoms related to advanced kidney cancer disease and is the sum of nine individual scores, HR=Hazard Ratio, ITT=intent-to-treat, NA=Not applicable, ne=Not estimable, NR=Not reported, OR=Objective response, OS=Overall Survival, PFS= Progression-free survival, PPE= palmar-plantar erythrodysaesthesia, TEAEs=Treatment-emergent adverse events, TTD=Time to deterioration.

[a] = adjusted for ECOG PS and prior treatment regimen.

[b] = TTD was assessed, where deterioration was defined as the composite endpoint of death or disease progression or a FKSI-15 decrease of ≥ 5 points, whichever occurred first.

[c] = ITT population.

[d] = ITT population censored for cross-over

[e] = Median PFS (by central review) for patients treated with everolimus versus placebo, after prior sunitinib (n = 184 patients, most of these received also prior sorafenib)[25]

[f] = Not corrected for cross-over/treatment switching. The rank-preserving structural failure time approach was also used to reconstruct the placebo survival curve as if all patients initially randomized to placebo never switched over to everolimus; the reconstructed median OS for placebo was 10.0 months, that is, 4.8 months shorter than the observed median OS with everolimus.

[g] = Results marked with a 'g' are from Motzer 2008 Lancet[26]; all other results for RECORD-1 from Motzer 2010 Cancer[25].

[h] = Rash or desquamation

[i] = Hand–foot skin reaction

[j] = Based on the final analyses of OS in November 2005, 6 months after cross-over (treatment switching) was allowed (216/452 had switched from placebo to sorafenib). HR for OS before cross-over: 072 (95% CI: 0.54, 0.94)

[k] = HR for PFS before cross-over: 0.44 (95% CI: 0.35, 0.55).

[l] = Based on ITT analyses of 903 patients in May 2005.
[m] = Results as reported in Rini 2011 Lancet[22].
[n] = Larger values are associated with better health states. A positive difference favours the first treatment.
[o] = As reported in additional files provided by the manufacturer in response to the clarification letter.
*) These data are from the Pfizer submission (MS, Table 21, page 101) and for the placebo arm include all patients having received sunitinib or sorafenib previously, some received up to 4 previous treatments. According to the Pfizer submission, 21% (n=89) of patients in RECORD-1, had received only one prior systematic therapy, sunitinib or sorafenib. The number of patients in the placebo arm having received only prior sunitinib is not known. According to Motzer et al. 2008[26], 124 out of 272 (46%) in the everolimus arm and 60 out of 138 (43%) in the placebo arm had received sunitinib only as prior systemic treatment.

4.2.2 Describe and critique the manufacturer's approach to validity assessment for each relevant trial.

Details of the critical appraisal of the AXIS, TARGET and RECORD-1 trials can be seen in Table 4.5 below.

Table 4.5: Critical Appraisal of included studies

Study question **Grade (yes/no/not clear/NA)**	**AXIS**	**TARGET**	**RECORD-1**
Was randomisation carried out appropriately?	Yes	Not clear[1]	Yes
Was the concealment of treatment allocation adequate?	Yes	Not clear[1]	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	No[2]	Yes[3]	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No	No	Not clear[4]
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	Yes

1 = Not sufficient details. 2 = Patients and investigators not blinded. PFS and ORR outcome assessors were blinded. 3 = It was a double blind study but disclosure was permitted after documentation of progression and the sorafenib patients who had response received open label treatment. 4 = There was some imbalance in drop outs between two groups. The placebo group had 77.7% drop outs while in the everolimus group 48% patients dropped out.

Comment

All three trials are good quality RCTs.

4.2.3 Describe and critique the statistical approach used within each relevant trial.

The statistical approaches used in the three trials, can be described as follows (based on the main publication for each trial):

AXIS:

The primary endpoint was PFS, defined as time from randomisation to either first documentation of RECIST defined disease progression (per independent radiology review of images) or death due to any cause, whichever came first. Secondary endpoints were overall survival, objective response rate, duration of response, and time to deterioration, a composite endpoint consisting of time to death, disease progression, or worsening of symptoms. Symptom deterioration was defined as two consecutive available decreases of at least five points from baseline using FKSI-15 (≥3 points using FKSI-DRS), unless it was the final score, for which one decrease was sufficient.

Efficacy in the intention-to-treat population was assessed on the basis of assessments by a blinded independent radiology review committee. This study was designed to test the hypothesis that treatment would result in an improvement in median PFS from five months with sorafenib, based on previous clinical data, to seven months with axitinib. The initial target sample size was calculated based on 90% power to show the improvement in PFS using a one-sided log-rank test at a significance level of 0.025. It was estimated that about 650 patients would need to be enrolled to observe 409 patients with progressive disease or death. The significance for efficacy analysis was calculated with the Lan-DeMets procedure with an O'Brien-Fleming stopping rule. For the secondary endpoint of overall survival, a total of 417 events were required for a log-rank test with an overall one-sided significance level of 0.025 to have power of 0.80, assuming a 31.67% improvement in median overall survival and a follow-up period of about 37 months. Kaplan-Meier methods were used to obtain estimates of median PFS. In accordance with the statistical design, a stratified, one-sided, log-rank test adjusting for ECOG performance status and previous treatment was used to compare PFS between the two treatment groups. Cox proportional-hazards models were used to explore potential effects of baseline stratification factors. A one-sided Cochran-Mantel-Haenszel test stratified by ECOG performance status and previous treatment was used to compare objective response rates between the two treatment groups. Duration of response was analysed using descriptive statistics. Symptom deterioration in the intention-to-treat population was assessed using the pre-specified time to deterioration composite endpoint. Survival analysis methods were used, including Kaplan-Meier plots and log-rank tests, to compare time to deterioration between the two treatment groups.[22]

TARGET:
The number of patients who would need to be enrolled was calculated in order to detect a 33.3% increase in overall survival among patients with sorafenib, as compared with those receiving placebo. Assuming a two-sided type I error of 0.04, the study would have 90% power to detect a 33.3% difference in survival between the two groups after a total of 540 patients had died. The duration of the study was estimated to be 29 months on the basis of the following assumptions: a monthly enrolment rate of 50 patients, an exponentially distributed event time, a median time of 12 months in the placebo group, and a 17-month long enrolment for a total of 856 patients in the two groups (428 per group). Assuming that 3% of patients would be lost to follow-up, approximately 884 patients had to be randomly assigned to study groups. According to these assumptions, approximately 270 deaths were expected in approximately 17 months.

Planned interim findings (when approximately 270 of the patients had died) and the final intention-to-treat findings regarding overall survival (when approximately 540 patients had died) were analysed with a stratified log-rank test. The O'Brien–Fleming spending function was used prospectively to ensure that the overall false positive rate (alpha) was no more than 0.04 (in a two-sided analysis). In the first analysis of overall survival, which was performed in May 2005, the information fraction — the total number of deaths (regardless of crossover) at the cut-off date divided by the total number of deaths specified by the protocol (540) — was used to calculate the O'Brien–Fleming threshold for significance (P = 0.0005).

In November 2005 (six months after cross-over was allowed), when the second analysis of overall survival was performed, the O'Brien–Fleming threshold was P = 0.0094. The final, planned analysis of overall survival was undertaken after 540 patients had died.

The planned, independently reviewed analysis of progression-free survival was performed on 28 January 2005, after disease had progressed in approximately 363 patients. The analysis had a power

of 90% to detect a 50% increase in progression-free survival in the sorafenib group (two-sided alpha of 0.01). Progression-free survival was compared by the log-rank test (stratified by prognostic group and country). All patients in the study groups were included in the efficacy analyses.

Treatment-related differences in response were evaluated by the Cochran–Mantel–Haenszel test. All patients receiving at least one dose of sorafenib were eligible for the safety analysis. All reported P values are two-sided and unadjusted for interim analyses.[23]

RECORD-1:

The sample size was calculated on the basis of the primary endpoint. A clinically meaningful improvement was defined as a 33% risk reduction (hazard ratio 0.67), corresponding to a 50% prolongation in median progression-free survival, from 3.0 months for the placebo arm to 4.5 months for patients receiving everolimus. With the two to one randomisation and assuming a one-sided cumulative α of 0.025, it was calculated that a total of 290 events as per central radiology review were required to achieve 90% power for the three-look group sequential plan. With a scheduled recruitment period of 16 months and additional follow-up of five months, it was estimated that about 362 patients would need to be enrolled (assuming that around 10% of patients would be lost to follow-up) to observe the required number of events.

The first and second interim analyses were planned after observing about 30% and 60%, respectively, of the targeted 290 events required for the final statistical analysis. These interim analyses allowed the study to be stopped on the basis of safety, or futility or efficacy (second analysis only). The final analysis was to be done when 290 progression events had been observed, if the stopping rule had not been met at an interim analysis.

After the second interim analysis, the study steering committee, on the recommendation of the independent data monitoring committee, decided to terminate the trial early because the pre-specified efficacy stopping boundary ($p \leq 0.0057$, determined according to the method of Lan-DeMets with O'Brien-Fleming-type stopping rules) was crossed, the null hypothesis rejected, and the criteria for a positive study met. This second interim analysis was designed to have 45% probability of detecting an effective treatment under protocol assumptions on the treatment effect. As per protocol, this second interim analysis was planned after observing about 60% of the targeted 290 progression-free survival events (per central radiology); however, because this central assessment was not done in real time and the number of events needed was unknown, the cut-off date (15 October 2007) was determined using a statistical prediction model based on events per the investigator.

The actual number of centrally assessed progression-free survival events observed as of the cut-off date and included in the analysis was 191 (or 66% of the targeted 290 events).

Patients without tumour progression or death at the time of the data cut-off for the analysis or at the time of receiving an additional anticancer therapy were censored at their last date of adequate tumour evaluation.

Progression-free and overall survival curves were estimated with Kaplan-Meier methodology; treatment arms were compared with a stratified log-rank test adjusting for strata defined by MSKCC prognostic score and the hazard ratio estimated by use of a stratified Cox proportional hazards model.[26]

Comment

The approach to the statistical analysis of trial data in the three trials was generally sound. The use of one-sided p-values in the AXIS trial is unusual but as they used a one-sided 0.025 level this is equivalent to the standard two-sided 0.05 significance level. The only contentious issue relates to attempts after the main trial analysis to adjust for placebo patients who switched to the active treatment (sorafenib or everolimus) in the TARGET and RECORD-1 trials after they had progressed. In TARGET the primary end point was overall survival, after an interim analysis revealed a significant benefit to sorafenib for progression-free survival the protocol was amended to allow placebo patients to cross-over to sorafenib. Overall 48% of patients crossed-over. An analysis which censored these patients at the point of cross-over was reported but this may be biased, more appropriate analysis methods (RPSFT models) were not reported in the trial. The RECORD-1 trial did however use more appropriate methods (RPSFT models) to adjust overall survival estimates for the time that the patient switched treatments. This issue has been discussed in detail and adequately addressed by the manufacturer in the submission (*see Pfizer submission, Section 6.7.2, page 79 (TARGET) and page 81 (RECORD-1*)).

4.2.4 Describe and critique the manufacturer's approach to outcome selection within each relevant trial.

The primary outcome measure in the AXIS trial was progression-free survival (PFS) as assessed by the independent review committee (IRC). The secondary outcome measures included in the trial were: PFS (Investigator assessment), overall survival, objective response rate (IRC and Investigator assessed), duration of response (IRC and Investigator assessed), patient reported outcomes (FKSI, FKSI-DRS, EQ-5D and composite endpoint time-to-deterioration (TTD)) and safety (*Source: Pfizer submission, Section 6.3.2, Table 5, page 45-48*).

The manufacturer uses outcome measures in accordance with those used in the AXIS trial[22] which concurs with the outcome measures specified in the final scope.[1] However, these outcomes are only used for the comparison between axitinib and sorafenib. For the comparison specified in the scope: axitinib versus BSC, only overall survival and PFS are presented.

4.2.5 To what extent does each relevant trial include the patient population(s), intervention(s), comparator(s) and outcomes as defined in the final scope?

Population: The population described in the scope is: "Adults with advanced renal cell carcinoma who have received prior systemic treatment". In the manufacturer's submission this has been modified to "Adult patients with advanced renal cell carcinoma (RCC) after failure of prior treatment with sunitinib or a cytokine", which is in line with the licensed indication (*see MS, Section 5, page 31*).

In the AXIS trial the population is described as: "Patients with mRCC following failure of a prior systemic first-line regimen containing one of the following: sunitinib, bevacizumab + IFNα, temsirolimus or cytokine(s)." In total, 126 out of 361 (35%) patients in the axitinib arm and 125 out of 362 (35%) patients in the sorafenib arm received cytokines as their one-and-only first-line treatment. While, 194 out of 361 (54%) patients in the axitinib arm and 195 out of 362 (54%) patients in the sorafenib arm received sunitinib as their one-and-only first-line treatment.

In the TARGET trial the population is described as: "Patients with metastatic clear cell RCC who had progressed after one prior systemic therapy in the previous 8 months". In total, 374 out of 451 (83%) patients in the sorafenib arm and 368 out of 452 (81%) patients in the placebo arm received cytokines as their one-and-only first-line treatment.

In the RECORD-1 trial the population is described as: "Patients with metastatic clear cell RCC who had progressed after sunitinib or sorafenib, or both. Previous therapy with bevacizumab, interleukin 2, or interferon alfa was also permitted." In total, 127 out of 272 (47%) patients in the everolimus arm received sunitinib as previous treatment, possibly in combination with sorafenib, bevacizumab, interleukin 2, and/or interferon alfa. The number of patients in the placebo arm having received (only) prior sunitinib is not known. According to the Pfizer submission, 21% (n=89) of patients in RECORD-1, had received only one prior systematic therapy, sunitinib or sorafenib (*MS, Section 6.7.11, page 98*). According to Motzer et al. 2008, 124 out of 272 (46%) in the everolimus arm and 60 out of 138 (43%) in the placebo arm had received sunitinib only as prior systemic treatment.[26]

Intervention: The intervention in the AXIS trial is axitinib, this is in accordance with the NICE scope. In the AXIS trial the intervention is described as follows:

> *"Axitinib was given orally at a starting dose of 5mg twice daily. Patients who tolerated the starting dose with no adverse reactions above grade 2 of Common Terminology Criteria for Adverse Events (CTCAE) for at least 2 weeks were allowed to have their dose increased, at the discretion of the treating physician, to 7mg twice daily, unless the patient's blood pressure was higher than 150/90 mm Hg or the patient was receiving anti hypertensive medication. Subsequently, with the same criteria, patients who tolerated the axitinib dose of 7mg had their dose increased to a maximum of 10mg twice daily. The axitinib dose could be reduced to 3 mg twice daily and then further to 2mg twice daily, if needed." (Source: Rini 2011, page 1932-33)*[22]

Comparators: The comparator specified in the NICE scope is best supportive care (BSC). There are no head-to-head comparisons of axitinib versus placebo/BSC. As described in section 3.3, the manufacturer's submission uses placebo as a proxy for BSC, which seems reasonable.

Outcomes: The three main trials (AXIS, TARGET and RECORD-1) use outcome measures in accordance with those specified in the final scope. However, relevant outcomes are often missing for the populations of interest (cytokine refractory and sunitinib refractory populations) and for the comparison axitinib versus BSC, we rely on an indirect comparison (cytokine refractory subgroup) and a simulated treatment comparison (sunitinib refractory subgroup). For these analyses, overall survival and PFS are the only outcomes presented.

Comment

The intervention and comparator are in accordance with the NICE scope. The main problem with the population is that there are no data from the RECORD-1 trial for patients who received sunitinib as their only first-line treatment.

4.2.6 Where appropriate, describe and critique any meta-analysis, indirect comparisons and/ or mixed treatment analysis carried out by the manufacturer.

Indirect Comparisons

Indirect comparisons were performed by the manufacturer. Two main analyses were performed, both using the same statistical method. These were: progression-free survival (PFS) for the cytokine refractory subgroup using the AXIS and TARGET trial data and overall survival (OS) for the cytokine refractory subgroup using the AXIS data and TARGET data which was censored for patients crossing from placebo to sorafenib. A further analysis of OS using database data was used in the simulated treatment comparison (details are reported in that section).

The methods used for the indirect comparisons were:

The indirect comparison was performed using Bayesian Markov Chain Monte-Carlo sampling to determine the relative efficacy of the treatments. Sampling was performed using WinBUGS. A fixed effects model was used due to the limited availability of relevant data for use in the model. In this case because hazard ratios entered to the model and not individual treatment effects, the approach assumes that the relative treatment effect (i.e. HR) for one treatment pair is the same across all trials. Since there was only one trial per pairwise HR, this assumption was appropriate in this analysis. Non-informative prior distributions were used. A non-informative prior assumes that all possible The WinBUGS code for the fixed-effects model is provided in Section 10.14 (Appendix 14 of the Manufacturer's Submission).

Point estimates of the HR for each pair of treatments along with 95% credible intervals (CrI) were calculated from 5,000 simulated draws from the posterior distribution after a burn-in of 20,000 iterations. (MS, Section 6.7.5, page 94)

The data used in the indirect comparison are presented in table 4.6.

Table 4.6: Input data

	AXIS (axitinib vs sorafenib)	**TARGET (sorafenib vs placebo)**
	HR (95% CI)	
PFS (IRC)	***Cytokine refractory population:*** 0.464 (0.318-0.676)	***ITT population:*** 0.44 (0.35-0.55)
OS	***Cytokine refractory population:*** 0.813 (0.555-1.191)	***ITT population censored for cross-over:*** 0.78 (0.62-0.97)

Abbreviations: CI, confidence interval; HR, hazard ratio; IRC, independent review committee; ITT, intent-to-treat; OS, overall survival; PFS, progression-free survival.

The results are presented in table 4.7.

Table 4.7: Results of the indirect comparison in the cytokine refractory population

Treatment comparison	**Median HR**	**95% CrI**
PFS Axitinib vs placebo	0.251	0.165-0.379
PFS Axitinib vs placebo*	0.203*	0.132-0.318*
OS Axitinib vs placebo	0.63	0.41-0.99

* Corrected analysis by ERG.

Comment

The code and data used in the mixed treatment comparisons were checked by the ERG group. An error was found in the trial data used in one of the analyses (progression free survival in the prior cytokine group) but the other two analyses were correct (overall survival in the prior cytokine group, and overall survival in the prior sunitinib group). Given the small number of trials included in the analyses the use of fixed effect models was appropriate.

The indirect comparison for progression free survival used the data reported in table 17 on page 93 of the manufacturer's submission. This gives a HR of 0.44 (0.35 to 0.55) for PFS from the TARGET trial. This gives a ln HR and SE of -0.82098055 (0.1153). However, the data used in the manufacturer's model for this trial was -0.616186139 (0.090) which is ln 0.54. When the model was

rerun the estimated median HR (95% CrI) was 0.203 (0.132 to 0.318) which is slightly smaller (and more favourable for axitinib) than 0.251 (0.165 to 0.379) as reported in table 18 on page 94 of the manufacturer's submission.

Simulated treatment comparison

As there was no direct or indirect trial evidence the manufacturer performed a simulated treatment comparison (STC). This compared PFS and OS for axitinib vs. everolimus and best standard care using data from the AXIS and RECORD-1 trials.

Simulated treatment comparison

Simulated treatment comparisons (STCs[27] is a novel technique to derive indirect comparisons between competing treatments (say A and B). Unlike mixed treatment comparisons (MTCs) which provide an average measure of the difference between A and B across all studies, STCs aim to answer a more specific question: what difference could we expect if A and B had been compared in the same trial.

STCs rely on individual patient data (IPD) for the treatment from an index trial (e.g., one used as the basis of a submission), and summary data (usually published reports) for the competitor from one or more studies. The studies for the treatment being compared must be generally compatible in terms of the type of population included, measurement methods, timeframe of observation, reporting of information, etc. The studies are not required to be exactly identical in these dimensions, but there must be sufficient overlap so that findings from one study can be assumed to be applicable in the setting of the other trial.

Even with close compatibility between the studies, it is unlikely that the characteristics of the patients will be identical, so that comparisons of outcomes between the trials may be confounded by these differences. STCs are specifically designed to adjust for these differences. This is done by using the index trial data to build a predictive equation for each endpoint for which a comparison is desired. We can denote this equation in a general way as having the following form:

$$\mu = X\beta$$

where μ represents some parameterization of the outcome variable. For instance, if the outcome of interest is a time-to-event variable like PFS or OS, μ would be the scale parameter in a parametric survival model; X represents a vector of predictors of the outcome and β represents the corresponding coefficients. We note that X may include an indicator for study group, and correspondingly, β would include a treatment effect coefficient. In some applications, the equation may be built from a single (e.g., experimental) treatment; in oncology trials, this may be done when outcomes in the reference arm is biased due to crossover, for example. For the explanations that follow, it is assumed that the equation is built from the primary treatment arm (i.e., A in the current notation).

The STC then proceeds with following steps:

1. *If the comparator treatment (B) had been included in the index trial the equation would have included a term for a comparison of A vs. B, as follows:*

$$\mu = X\beta + \delta_{B\ vs.\ A} Z_{B\ vs.\ A}$$

where δ is a coefficient representing the effect of B compared to A (e.g., expressed as a log hazard ratio), and Z is an indicator of treatment group.

2. *Since the index trial provides no information on treatment B, external data from published sources must be used to estimate δ. For time-to-event outcomes like PFS or OS, this information may be in the form of a Kaplan-Meier curve or specific percentiles of the time-to-event distribution, like the median.*

3. *δ can then be estimated by calibrating the equation to the target values in step 2; that is, finding a value that will yield a predicted outcome that equal the target values (e.g., median survival) established in step 2. To account for the fact that this target value reflects outcomes in the population for treatment B, the predictions must be adjusted to the profile of the comparator's study.*

   This is done by setting X to the mean characteristics of the population in study B:

   $$\mu_B = X_B\beta + \delta_{B\ vs.\ A} Z_{B\ vs.\ A}$$

   This represents outcomes for patients like those in the competitor study, had they received treatment A (since $X\beta$ predicts outcomes for treatment A). Thus, the difference between predictions based on $X_B\beta$ (e.g., the median time) and the target value (e.g., median time observed in study B) reflects the difference in the effects of treatment A and B.

4. *The value of δ is then a function of this difference in outcome measures. This may be calculated algebraically in situations where the target values are simple numeric values (e.g., medians). When the target is a distribution (e.g., Kaplan-Meier curve), a grid search may be performed to identify the value that minimizes differences between the prediction and target values.*

   Predictive Equations for PFS and OS:

   Patient level data from the AXIS trial were analysed to derive parametric failure-time (survival) equations incorporating baseline predictors of the endpoint. These equations were based on the axitinib arm only. Patient-level data were analysed using exponential, Weibull, Gompertz, lognormal and loglogistic distributions (using Stata 10.0). Data were fitted to the clinical survival data for the axitinib treatment arm separately for the cytokine refractory and sunitinib refractory subgroups. Model fit was assessed using Akaike's information criteria (AIC) and Bayesian information criteria (BIC), smaller values indicate a better fitting model, as well as visual inspection of fitted survival curves against Kaplan-Meier curves. Of the five distributions examined in the full parametric survival analysis, the two best fitting (log-normal and Weibull) were used in the STC.

   Identification of potential outcome predictors:

   From the AXIS patient level data and prior clinical knowledge, predictive factors were identified that may have been influential on the length of the final PFS or OS. These included sex, age, nephrectomy status, previous radiotherapy, previous cytokine therapy, MSKCC score, clear cell carcinoma, ECOG performance status and time on sunitinib treatment. Univariate regression analyses were performed to determine which of these factors were predictive of PFS and /or OS. That is, one factor at a time was analysed to determine which resulted in significantly longer/ shorter PFS or OS and these were included in a multivariate equation (one for PFS and one for OS). Characteristics that were identified as being predictive in the univariate analyses (i.e. having a statistically significant coefficient with p-value <0.10) were then considered further. Multivariate analyses incorporated these characteristics simultaneously and the final equations were determined by manually trimming the model to include only significant predictors (p values <0.10).

Validating the equations:

The final equations were checked for validity, i.e. that they aligned with clinical knowledge, and their ability to replicate the source data. These equations formed the basis for the simulation of the "missing arms".

Target Values for Comparisons of Axitinib vs. Everolimus and BSC

Ideally, the STC would rely on calibration to the full observed Kaplan-Meier survival curve for everolimus and BSC, but these were not reported for the relevant RECORD-1 populations. Therefore, calibration was carried out using the median PFS and OS times. Calibration to the median assumes that everolimus, BSC and axitinib curves for OS and PFS arise from the same type of survival distribution with a common shape.

Since survival estimates for the prior sunitinib placebo only population were not reported for the RECORD-1 study, two data sources were examined for the comparison, each necessitating different assumptions:

***ITT RECORD-1 placebo**: As the prior sunitinib placebo population was not available, the first approach taken was to compare the AXIS sunitinib refractory patients with the ITT placebo population of RECORD-1. As the RECORD-1 ITT placebo population includes patients that have previously received sunitinib and/or sorafenib, this approach assumes that prior sunitinib patients have equivalent patient characteristics and outcomes to prior sorafenib patients. The median PFS and OS estimates of this patient population are 7.8 weeks (1.8 months), and 43.4 weeks (10.0 months), respectively. Due to cross-over in the RECORD-1 trial, median reported OS for BSC group from RPSFT analysis (i.e. 10 months) was used for calibration of the OS curve. However median OS of 10 months was from the RPSFT analyses using the entire BSC cohort and not sunitinib-refractory patients only, therefore the adjustment factor derived from this analysis is likely to be conservative. This is supported by evidence from the RECORD-1 study where prior sunitinib patients receiving everolimus had median OS of 12.6 months compared to 14.8 months in the ITT population.*

***RECORD-1 prior sunitinib everolimus**: The second approach taken was to compare the prior-sunitinib AXIS patients to the prior sunitinib RECORD-1 patients in the everolimus treatment arm (denoted as prior sunitinib everolimus). This population was reported by DiLorenzo et al[28](see 98 MA 2011, EOP) and achieved median PFS and OS times of 16.9 (3.9 months) and 54.4 weeks (12.6 months), respectively. Median PFS for everolimus patients who failed prior sunitinib was taken from Motzer et.al, 2010[25] due to results presented in Di Lorenzo contradicting Motzer et al (i.e., 5.6 months vs. 3.9 months median PFS for sunitinib-refractory patients). An attempt was made to follow up with the authors to clarify the discrepancy in these two measurements, however, it is still unclear how the results in the Di Lorenzo study were obtained or why they contradict the previous publication.*

Since these patients received everolimus, the survival curves generated by the STC were required to be further adjusted by the application of the PFS and OS hazard ratios from the RECORD-1 study (between everolimus and placebo) to create modelled "AXIS-like" placebo curves. This was done by applying the hazard ratio from the RECORD-1 study to the STC curve after the STC was completed. This approach does not require the assumption of similar characteristics and outcomes between the RECORD-1 prior sunitinib and ITT population. However, as the hazard ratios used to model the everolimus-placebo PFS and OS relationships are from the AXIS ITT population, it does require the assumption of equivalent incremental efficacy for everolimus vs BSC between the prior sunitinib and RECORD-1 ITT population.

As neither one of these assumptions was considered de facto more valid than the other, the STC explored both approaches. (MS, Section 6.7.11, page 98-101 and section 7.3.2.1 page 138)

To support the STC a further analysis using data from a Swedish database was performed.

Due to the lack of published work reporting on the survival of patients that progressed on first-line sunitinib treatment and then received BSC, a retrospective analysis of sunitinib-refractory patients from a Swedish database (Renal Comparison; RENCOMP) containing data from three registries (The Swedish Cancer register, The National Patient Register and The Swedish Prescribed Drug Register) was carried out to determine the OS of patients who received sunitinib first-line, followed by BSC or sorafenib second-line.

The aim of this comparison was to estimate the OS hazard ratio between patients who received sunitinib followed by sorafenib and sunitinib followed by BSC. These estimated hazard ratios using RENCOMP were then included in an indirect comparison alongside the AXIS sunitinib refractory hazard ratio between axitinib and sorafenib to generate indirect hazard ratios between axitinib and BSC in a sunitinib refractory population (MS, page 83-84).

The study utilised in this submission is a sub-analysis of a larger retrospective, non-interventional study carried out using data collected and stored in three comprehensive linked registries by the National Board of Health and Welfare, Stockholm, Sweden (see table 4.8 below for a description of the registries included). This study, known as the the RENal COMParison (RENCOMP) study, has been previously published.

Table 4.8: Summary of National Swedish Registries used in the RENCOMP study

Registry	Year Founded	Data	% of population covered
Swedish Cancer Register	1958	Diagnosis and death records for all patients with a cancer diagnosis	100
National Patient Register	1987	Information on inpatient visits (since 1987) and outpatient visits (since 2001)	>90
Swedish Prescribed Drug Register	2005	Dates and amounts of prescribed and dispensed drugs for individual patients	100

To estimate the relative efficacy of sorafenib vs. BSC on overall survival, this study examines real-world retrospective data to compare the OS of patients who received either sunitinib followed by sorafenib with those who received sunitinib followed by BSC. The current analysis includes 135 patients who were identified with advanced/mRCC and were recorded as having received first-line treatment with sunitinib after the introduction of TKIs in Sweden in 2006.

In order to correct for confounding factors (i.e. patient characteristics that may have been different between the two treatment arms in the database), a multivariate Cox proportional regression analysis was performed to create adjusted hazard ratios for sorafenib vs BSC in the second-line setting.

Covariates tested in the model were aligned with those included in two previous RENCOMP publications[29, 30] with several additional covariates included based on alignment with known

mRCC prognostic factors typically included in clinical trials. The regression model included the following covariates:

Lead Time for Diagnosis: *In accordance with Motzer criteria for mRCC, a dummy variable for the lead time between RCC diagnoses and mRCC was constructed and denoted 'Lead time RCC-met (1 year +, vs < 1year). A longer interval between RCC diagnosis and metastatic disease would indicate healthier patients and imply a longer chance of survival.*

Age: *A dummy variable for age defined as 'Age_met_65' which was =1 if age was greater than 65 at the start of second-line treatment, and =0 if age was 65 years or less at the start of treatment. A higher age would imply a lower OS.*

Lead Time for Treatment: *There is a wait and see tradition in mRCC treatment for patients that have a good prognosis (e.g. indolent disease, minimal metastatic sites, good performance status). Therefore, the variable 'Leadtime_mRCC_firstpre' was constructed with a value = 1 if lead time was less than 1 year and = 0 if lead time was 1 year or longer. A shorter lead time would hence indicate sicker patients with lower survival chances.*

Duration of Sunitinib Treatment: *A longer duration of sunitinib treatment may indicate stronger likelihood of survival in the second-line setting, as demonstrated in the patient level data analysis of the AXIS sunitinib refractory patients (see Section 6.7.11-Manufacturer's submission). Therefore a dummy variable was constructed to account for this, defined as 'Days of SU treatment' = 1 if duration was 90 days (3 months) or more and < 1 if duration was less than 3 months.*

The results and explanatory power of the analysis may be affected by the number of variables included in the model. The choice of variables incorporated in the base-case model was aligned with variables reported as significantly affecting OS in the previous RENCOMP publications. Sensitivity analysis for different combinations of explanatory variables was carried out to examine the model for robustness

Details of the patient characteristics from the RENCOMP data are presented in table 4.9.

Table 4.9: Patient characteristics – RENCOMP (*Source: Pfizer submission section 6.7.11, page 101*)

	Sorafenib N=59	**BSC N=76**
Male, %	72.9	69.7
Nephrectomy, %	79.7	75.0
>65 years of age at second-line treatment, %	62.7	53.9
RCC diagnosed 2000-2005, %	14.8	35.7
RCC diagnosed 2006-2008, %	85.2	64.3
mRCC diagnosed 2000-2005, %	5.1	9.9
mRCC diagnosed 2006-2009, %	89.8	86.8
Days_since_RCC_met < 1 year, %	64.4	56.6
M1 at diagnosis, %	35.6	26.3
Leadtime_mRCC_firstpre_ <1 year, %	83.1	75.0
>90 days sunitinib treatment, %	84.7	56.5
Treated at a large institution, %	33.9	40.8
Region, %		
South region	25.4	34.2
Mid Central Region	6.8	6.6
Stockholm Region	27.1	25.0
East Region	3.4	5.3
North Region	8.5	13.2
West Region	28.8	15.8

Abbreviations: BSC, best supportive care; mRCC, metastatic renal cell carcinoma; RCC, renal cell carcinoma.

The results from the analysis of the RENCOMP data are summarised in table 4.10 below.

Table 4.10: Multivariate Cox proportional-hazards regression analysis of RENCOMP data

	Base case	
	Hazard ratio (95% CI)	**P value**
Second-line treatment (sorafenib vs BSC)	0.621 (0,412, 0,936)	0.023
Age 2nd line treatment start (age ≥ 65 vs <65)	0.754 (0.496, 1.144)	0.754
Gender (female vs male)	0.747 (0.460, 1.213)	0.239
Nephrectomy (yes vs no)	0.509 (0.317, 0.817)	0.005
Lead time between RCC and mRCC (≥ 1 year vs <1 year	0.629 (0.405, 0.979)	0.040

Abbreviations: BSC best supportive care; CI confidence interval; mRCC, metastatic renal cell carcinoma; RCC, renal cell carcinoma.

This hazard ratio for sorafenib versus best supportive care (BSC) was used in an indirect comparison with the sunitinib refractory OS hazard ratio from the AXIS study (0.997, 95% CI 0.782, 1.27) to estimate an axitinib-BSC OS hazard ratio for patients who had received prior sunitinib of 0.619 (95% CrI 0.384 to 0.997).

STC results for progression-free survival

The results showing significant predictors of PFS from the accelerated failure time survival models are shown in table 4.11 below:

Table 4.11: Predictors of PFS and associated coefficient estimates

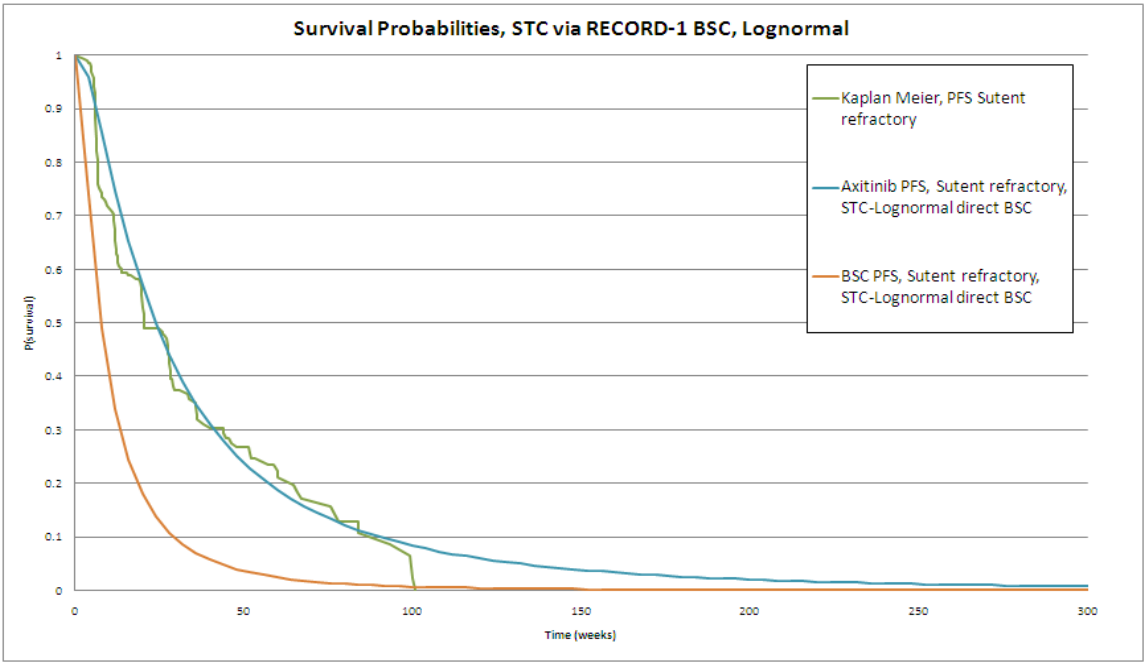
Predictors	**Lognormal estimate [95% CI]**	**Weibull estimate [95% CI]**
Intercept	0.5455 (-0.3277;1.4186)	0.8065 (-0.0339;1.6468)
MSKCC		
Favourable vs poor/NA	0.8405 (0.4116;1.2695)	0.8575 (0.4352;1.2799)
Intermediate vs poor/NA	0.241 (-0.0928;0.5747)	0.2256 (-0.0896;0.5409)
Age	0.0149 (0.0009;0.0289)	0.0179 (0.0038;0.032)

Abbreviations: CI, confidence interval; MSKCC, Memorial Sloan-Kettering Cancer Centre; N/A, not available; PFS, progression-free survival.

The best fitting distributions for use in the STC were the lognormal and Weibull. For the ITT RECORD-1 placebo cohort, the derived adjustment factor from the lognormal distribution was -1.12, corresponding to a median of 6.9 weeks (1.6 months) for axitinib-like patients if they had received placebo. The derived adjustment factor from the Weibull distribution was -1.25 (HR=4.0 for placebo vs. axitinib), corresponding to a predicted median of 7.4 weeks (1.7 months).

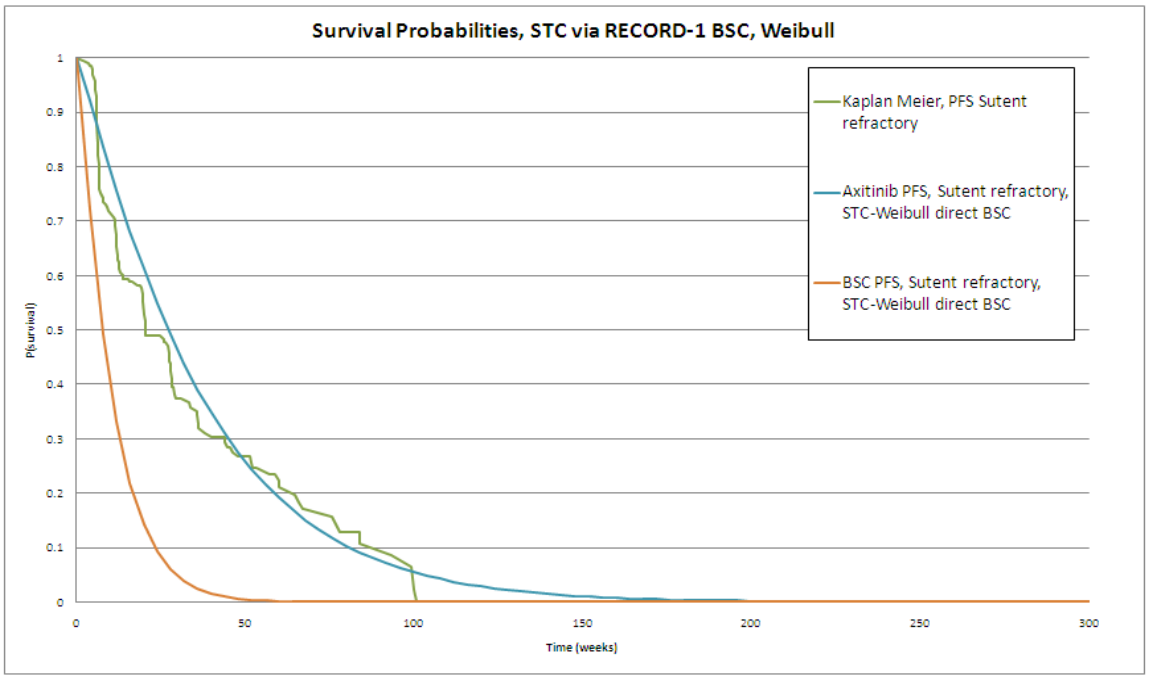
Figure 4.2 and 4.3 display the survival probabilities for the lognormal and Weibull curves, respectively.

Figure 4.2: Lognormal PFS distribution via RECORD-1 ITT placebo population *(Source: Pfizer Submission, page 102)*



Abbreviations: BSC, best supportive care; PFS, progression-free survival; STC, simulated treatment comparison.

Figure 4.3: Weibull PFS distribution via RECORD-1 ITT placebo population *(Source: Pfizer Submission, page 103)*



Abbreviations: BSC, best supportive care; PFS, progression-free survival; STC, simulated treatment comparison.

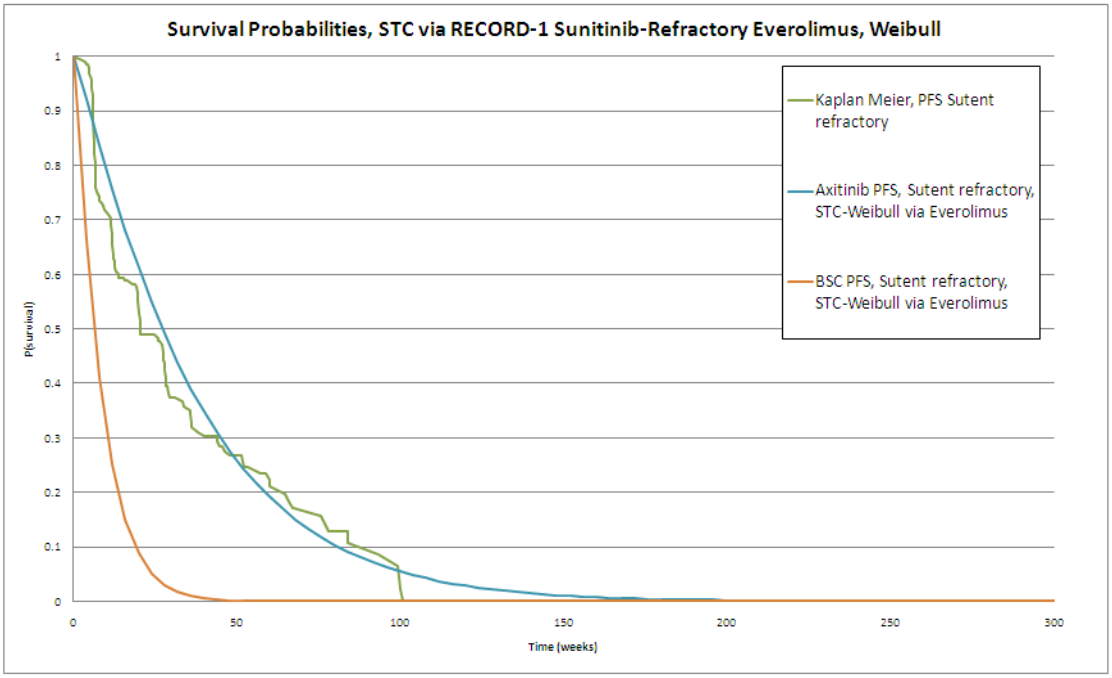
Calibrated PFS for axitinib-like patients - RECORD-1 prior sunitinib everolimus patients

For the RECORD-prior sunitinib everolimus patients, the derived adjustment factor from the lognormal distribution was -0.35, corresponding to a median PFS of 15.6 weeks (3.6 months). The derived adjustment factor from the Weibull distribution was -0.47, corresponding to a predicted median of 15.7 weeks (3.6 months).

The prior sunitinib PFS hazard ratio from the RECORD-1 study (HR =0.34; 95% CI: 0.23-0.51) was applied to the everolimus STC curve to generate a modelled AXIS-like, prior sunitinib PFS curve. As the lognormal distribution does not support the application of a hazard ratio, the Weibull was the only option explored in the model.

Figure 4.4 displays the survival probabilities from the Weibull distribution.

Figure 4.4: Weibull PFS distribution via RECORD-1 prior sunitinib everolimus + prior sunitinib HR if both treatments had been included in AXIS RCT for sunitinib-refractory patients
(Source: Pfizer Submission, page 104)



Abbreviations: BSC, best supportive care; PFS, progression-free survival; STC, simulated treatment comparison.

Table 4.12 presents a summary of the STC results for PFS for the sunitinib-refractory population.

Table 4.12: Predicted (Mean and Median) PFS times from STC

	Observed median PFS (months)	Predicted median with Weibull (months)	Predicted median with Lognormal (months)	Axitinib vs. placebo Estimated difference in mean PFS (months) Weibull/ Lognormal)
RECORD-1 ITT placebo population	1.8	1.6	1.7	[redacted]
RECORD-1 Prior sunitinib everolimus population	3.9	3.6	3.6	[redacted]

Abbreviations: BSC, best supportive care; TT, intent-to-treat.

STC results for overall survival

The results showing significant predictors of OS from the accelerated failure time survival models are shown in table 4.13 below:

Table 4.13: Predictors of OS and associated coefficient estimates (*source Pfizer manuscript page 105*)

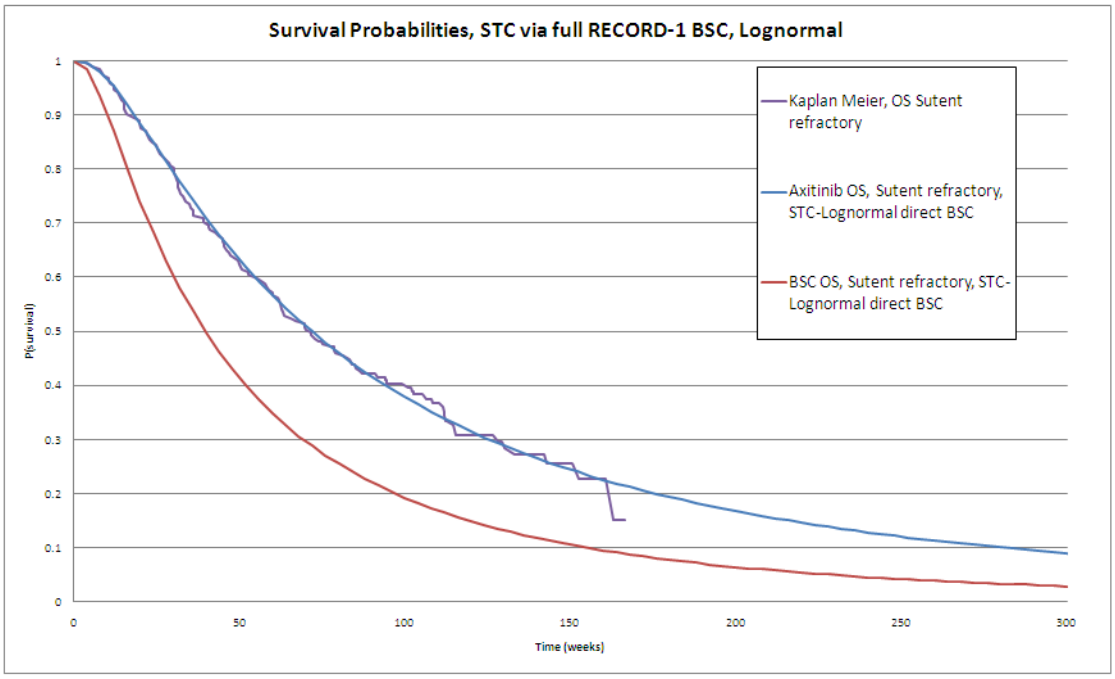
Predictors	Lognormal estimate [95% CI]	Weibull estimate [95% CI]
Intercept	2.0956 (1.8166;2.3746)	2.625 (2.369;2.8809)
MSKCC		
Favourable vs poor/NA	1.5225 (1.0983;1.9467)	1.3968 (0.9084;1.8851)
Intermediate vs poor/NA	0.5983 (0.2981;0.8985)	0.4929 (0.2183;0.7675)
Duration of prior sunitinib	0.0029 (-0.0005;0.0064)	0.0013 (-0.0021;0.0046)

Abbreviations: CI, confidence interval; MSKCC, Memorial Sloan-Kettering Cancer Centre; N/A, not available; PFS, progression-free survival.

The best fitting distributions for use in the STC were the lognormal and Weibull. For the ITT RECORD-1 placebo cohort, the derived adjustment factor from the lognormal distribution was -0.59, corresponding to a median of 36 weeks (8.3 months) for axitinib-like patients if they had received placebo. The derived adjustment factor from the Weibull distribution was -0.68 (HR=2.33 for placebo vs. axitinib), corresponding to a predicted median of 35.6 weeks (8.2 months).

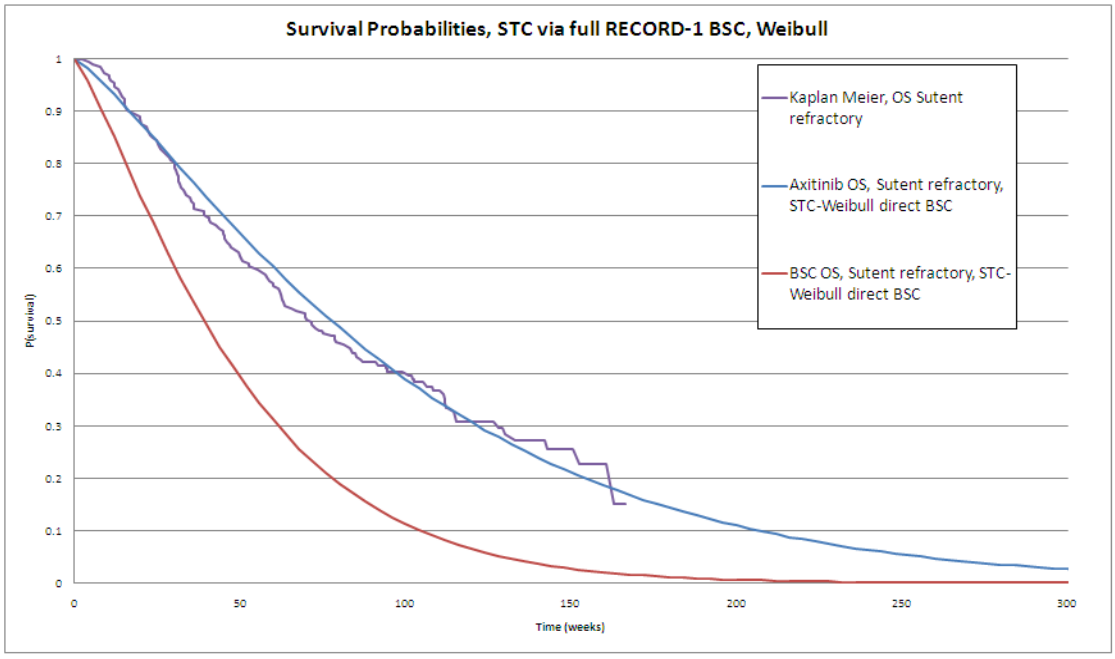
Figures 4.5 and 4.6 display the survival probabilities for the lognormal and Weibull curves, respectively.

Figure 4.5: Lognormal OS distribution via RECORD-1 ITT placebo population if both treatments had been included in AXIS *(Source: Pfizer Submission, page 105)*



Abbreviations: BSC, best supportive care; OS, overall survival; STC, simulated treatment comparison.

Figure 4.6: Weibull OS distribution via RECORD-1 ITT placebo population if both treatments had been included in AXIS RCT *(Source: Pfizer Submission, page 106)*



Abbreviations: BSC, best supportive care; OS, overall survival; STC, simulated treatment comparison.

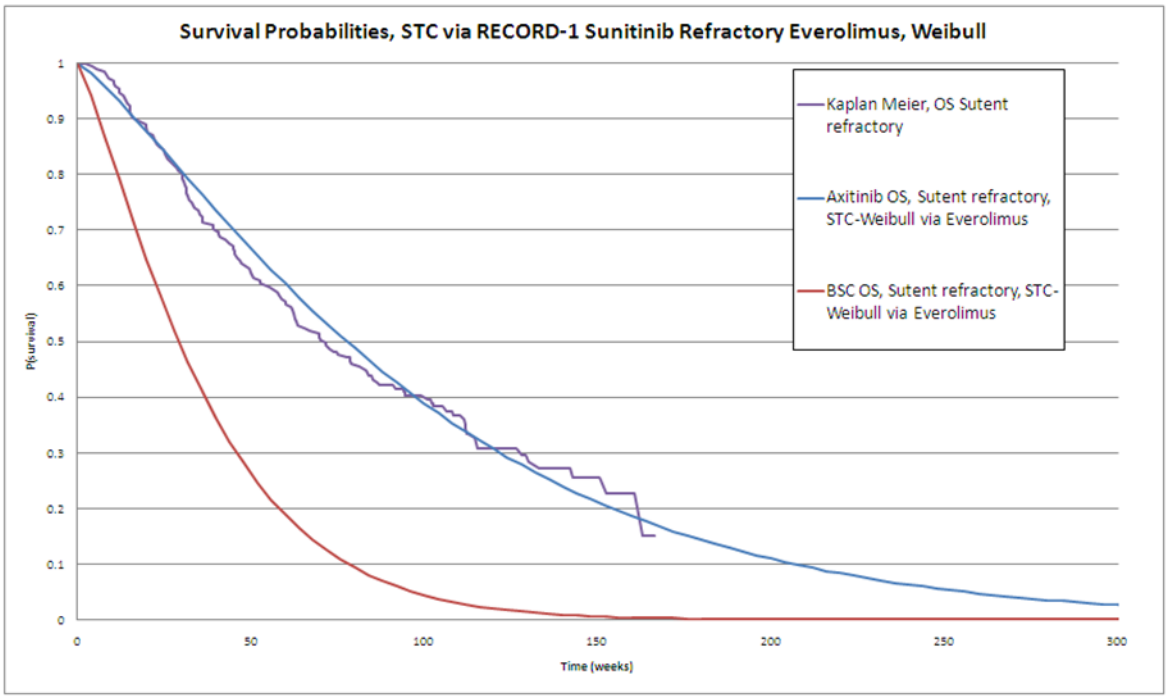
Calibrated OS for axitinib-like patients - RECORD-1 prior sunitinib everolimus patients

For the RECORD-1 prior sunitinib everolimus patients, the derived adjustment factor from the lognormal distribution was -0.37, corresponding to a median OS of 46 weeks (10.6 months). The derived adjustment factor from the Weibull distribution was -0.46, corresponding to a predicted median OS of 45.4 weeks (10.5 months).

The prior sunitinib PFS hazard ratio from the RECORD-1 study (HR =0.34; 95% CI: 0.23-0.51) was applied to the everolimus STC curve to generate a modelled AXIS-like, prior sunitinib PFS curve. As the lognormal distribution does not support the application of a hazard ratio, the Weibull was the only option explored in the model.

Figure 4.7 displays the survival probabilities from the Weibull distribution.

Figure 4.7: Weibull OS distribution via RECORD-1 prior sunitinib everolimus + RPSFT HR if both treatments had been included in AXIS RCT *(Source: Pfizer Submission, page 107)*



Abbreviations: BSC, best supportive care; OS, overall survival; STC, simulated treatment comparison.

To create a modelled placebo arm for the everolimus prior sunitinib population, the RPSFT-adjusted OS hazard ratio from the RECORD-1 study (0.53) was applied to the AXIS-like everolimus curve to generate a modelled AXIS-like, sunitinib refractory placebo curveAs the lognormal model does not support the application of a hazard ratio, the Weibull was the only option explored in the model. Figure 8 displays the survival probabilities calculated using the Weibull distribution.

Table 4.14 presents a summary of the STC results for OS for the sunitinib-refractory population.

Table 4.14: Predicted (Mean and Median) OS times from STC (*source Pfizer manuscript page 107*)

	Observed median OS (months)	**Predicted median OS with Weibull (months)**	**Predicted median OS with Lognormal (months)**	**Axitinib vs. placebo Estimated difference in mean OS (months) Weibull/ Lognormal)**
RECORD-1 ITT placebo	10.0	8.2	8.3	[redacted]
RECORD-1 Prior sunitinib everolimus	12.6	10.5	10.6	[redacted]

Abbreviations: BSC, best supportive care; TT, intent-to-treat.

Comment

It was not possible to check the results of the STC as this used the actual individual patient data from the manufacturer's trial. However the analysis seems to have been performed correctly and the reporting of the methods, results and limitations is clear. The STC used an advanced, but seemingly recent and rarely used, method to estimate the comparison between axitinib and placebo. This involved using the axitinib data from the AXIS trial to derive a predictive equation for overall survival based on patient baseline characteristics, calibrating this equation using external placebo data from another study (RECORD-1), and then using this new equation to predict survival results for placebo patients in the original trial (therefore replicating their results if there had been a placebo arm in the original trial). The main limitation is due to the assumption of comparability between the trials, but this assumption applies to all meta-analyses. However, this comparison is not based on randomised treatment allocation, but on a comparison of two single treatment arms; therefore there is considerable potential for bias in the outcomes of this analysis. Another limitation is due to the analysis including observational data from the RENCOMP database. Observational evidence is prone to more bias than evidence from RCTs and is considered a lower level of evidence. Combining RCT and observational data creates, as the manufacturer states, "a potential source of uncertainty" as the reasons for discontinuing the first-line treatment were not known so these patients may have different characteristics to those in the trial. The analysis of the RENCOMP data was also based on a small sample size (59 sorafenib and 76 for best supportive care). Finally the uncertainty around the STC results is unclear. There is no presentation of an associated SE or 95% CI for any of the results. This is a major omission as we have no idea of the uncertainty of the estimates. The estimates of mean or median PFS or OS and the associated difference between axitinib and placebo should all have been reported with associated variance estimates. The manufacturer's submission discusses the limitations in trial comparability and the impact on the results of their STC analyses on pages 108 to 109. This discussion seems reasonable and as the STC analysis used the only trial data available, this seems to be the only possible analysis which could have been performed under the circumstances.

4.2.7 Additional clinical work conducted by the ERG

We performed a more sensitive update search to try to bridge the gap between axitinib and BSC/placebo in the sunitinib refractory population (see section 4.1.2 of this ERG report). We found one study comparing temsirolimus with sorafenib in a sunitinib refractory population.[21] As this study is sponsored by Pfizer and the primary completion date is January 2012 Pfizer should be able to access data from this study. This does add another treatment that can be compared indirectly with axitinib. However, as we were not able to link temsirolimus with either everolimus or placebo in the relevant population, we were also not able to bridge the gap.

No additional analyses were conducted by the ERG in relation to clinical effectiveness.

4.3 Conclusions

There is no direct evidence for the comparison axitinib versus BSC. Instead the submission relies on an indirect comparison for cytokine refractory population, using evidence from two RCTS and a simulated treatment comparison for the sunitinib refractory population, using evidence from single treatment arms from two trials.

On page 128 of the MS the manufacturer justifies the choice of an active comparator in the AXIS trial, instead of BSC as requested by NICE, by stating:

> *"It was not considered ethical, with the availability of a licensed second-line medication, to provide patients with placebo. It was also considered that the use of an active comparator would provide a more robust analysis of the efficacy and safety of axitinib." (Source: Pfizer submission, Section 6.10.3, page 128)*

While this may be true for the cytokine refractory subgroup, it definitely is not true for the sunitinib refractory population. Sorafenib is approved in second-line for a cytokine refractory population only, reflective of the TARGET study, and not approved in a VEGFR-TKI refractory population. Therefore, for the sunitinib refractory population in the AXIS trial, BSC would have been the correct comparator at the time of the AXIS study start-up, as it is today. In fact, it would have been unethical to use sorafenib as the comparator in this population, as it was not approved for this population.

For the cytokine refractory population, the evidence relies on an indirect comparison, including 251 patients in the AXIS trial and 742 patients in the TARGET trial. Comparison of trial populations used in the analyses is not possible, because patient characteristics are not reported for cytokine refractory patients separately in both trials. When the full trial populations are compared the trials are reasonably comparable; the main difference being slightly better MSKCC-scores in the TARGET trial. In addition, the TARGET trial only reported liver and lung metastases, whilst AXIS reported a broader list of metastatic sites.

Another problem for this comparison is concern about the results for overall survival in the TARGET trial due to treatment switching from placebo to sorafenib at the point of disease progression.

For the sunitinib refractory population, the evidence relies on a simulated treatment comparison, including 194 patients in the axitinib-arm of the AXIS trial and 139 patients in the placebo-arm of the RECORD-1 trial. The placebo-arm of the RECORD-1 trial included mostly patients who had received more than one previous treatment and patients could have been treated with sunitinib or sorafenib, as well as a cytokine in many cases. The treatment arms are reasonably comparable on the patient characteristics reported, the main difference being slightly better ECOG/KPS and MSKCC-scores in the RECORD-1 placebo-arm. However, the fact that this comparison is not based on randomised treatment allocation, but on a comparison of two single treatment arms, means that there is considerable potential for bias in the outcomes of this analysis.

As reported by the manufacturer, the baseline characteristics for the prior sunitinib patients in the RECORD-1 study were not reported separately; therefore, only characteristics of the whole placebo population could be utilised in the comparison with the AXIS trial patient population *(See: Pfizer submission, Section 6.7.11, page 98).*

Other differences between AXIS and RECORD-1:

Despite some similarities between RECORD-1 and AXIS in terms of requiring at least one prior treatment, there are several differences between the two trials which could potentially confound the comparison. First, in contrast to AXIS, where all patients included in the study were required to have progressed on first-line therapy by RECIST-defined criteria, in the overall RECORD-1 population, 14% of patients (n=58) discontinued previous TKI therapy because of unacceptable toxicity. Among the subgroup of 58 patients who were intolerant to previous TKI therapy, 45 patients and 13 patients were randomly assigned to everolimus and placebo, respectively. Thus, patients in the RECORD-1 study could have discontinued prior treatment due to intolerance and therefore results would be more reflective of a first-line study.

Secondly, only 43 patients in the everolimus arm of RECORD-1 had sunitinib as there only previous therapy (i.e. purely second-line) in comparison with 194 patients in the AXIS trial. Of the 43 sunitinib refractory patients in RECORD-1, it was not known how many patients entered the trial due to sunitinib intolerance [31]. The inclusion of patients who were sunitinib intolerant rather than refractory would potentially bias the results in favour of the RECORD-1 patients; those patients who discontinue treatment due to intolerance can be considered to be analogous to first-line patients and would be expected to respond better to treatment compared with patients who failed first-line treatment.

Thirdly, in contrast to the AXIS study, where patients were required to have received only one prior therapy (sunitinib or a cytokine, or bevacuzimab + interferon-α or temsirolimus), patients in the RECORD-1 study were allowed to have received more than one previous therapy and could have been treated with sunitinib or sorafenib, as well as a cytokine in some cases. (Source: Pfizer submission, Section 6.7.2, page 81-82)

Another consideration in comparing the two studies is that RECORD-1 study patients were allowed to have received previous treatment with sorafenib as well as sunitinib. However, the median OS and patient characteristics have never been reported for those patients in the BSC arm that had progressed on sunitinib after receiving only one line of therapy. The closest available patient populations reporting overall survival data to allow the STC comparison were the ITT BSC population (corrected for crossover using the RPSFT method) and patients receiving everolimus treatment with only prior sunitinib therapy. (Source: Pfizer submission, Section 6.7.11, page 108)

Finally, the manufacturer tried to support the results of the STC by using data from a Swedish patient registry (RENCOMP):

Due to the lack of published work reporting on the survival of patients that progressed on first-line sunitinib treatment and then received BSC, a retrospective analysis of sunitinib-refractory patients from a Swedish database (Renal Comparison; RENCOMP) containing data from three registries (The Swedish Cancer register, The National Patient Register and The Swedish Prescribed Drug Register) was carried out to determine the OS of patients who received sunitinib first-line, followed by BSC or sorafenib second-line. (Source: Pfizer submission, Section 6.7.2, page 83-84)

This analysis is potentially biased as the patients are not randomly allocated to second-line treatment. There could be all sorts of underlying reasons why patients receive BSC or sorafenib following sunitinib that could influence treatment outcomes.

It is suggested in the MS that STCs have been accepted in previous appraisals by NICE (*Pfizer submission, Section 6.7.2, page 81*). In the manufacturer's response to the clarification letter, the

manufacturer clarified that an STC analysis was used and accepted by NICE in the appraisal of lenalidomide for multiple myeloma (NICE TA171[32]) (*Source: Manufacturer's response to clarification letter, Section A, Question 5*). The main difference between this submission and TA171 is that the statistical methods used in TA171 formed part of the economic model, not the meta-analysis of evidence from the systematic review as is the case in this submission. However, the statistical methodologies used appear to be similar; except TA171 does not refer to the method as a Simulated Treatment Comparison, this phrase does not feature in the document.

The manufacturer was "not aware of STCs being used in any other NICE appraisals" (*Source: Manufacturer's response to clarification letter, Section A, Question 5*).

5 COST EFFECTIVENESS

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

5.1.1 Objective of cost effectiveness review

The objective of the cost effectiveness review in the MS was to identify cost-effectiveness studies from the literature for the treatment of advanced/mRCC after failure of prior systemic treatment.
The search strategies for the cost-effectiveness review are discussed in detail in section 4.1.1.

5.1.2 Inclusion and exclusion criteria

Table 5.1 presents an overview of inclusion and exclusion criteria used for the review. The ERG considers the in- and exclusion criteria appropriate for the current review.

Table 5.1: Inclusion and exclusion criteria used for the review

Criteria	Include	Exclude
Population	Adults with advanced or metastatic RCC	Paediatric populations and other indications
Intervention	Axitinib	First-line treatments and non-pharmaceutical interventions
Comparator	• Pazopanib • Sunitinib • Sorafenib • Interferon-α • Interleukin-2 • Everolimus • Temsirolimus • Bevacizumab in combination with interferon • Tivozanib	Other interventions not licensed in RCC and combination therapies.
Outcomes	Cost outcomes (e.g. total costs, costs per life year gained, costs per QALY gained, ICER, ICUR)	NA
Setting	Any	Not limited

Study design	• Economic evaluations
	  ○ Cost-benefit analysis
	  ○ Cost-effectiveness analysis
	  ○ Cost-utility analysis
	  ○ Cost-minimisation analysis
	  ○ Cost-consequence analysis
		Cost studies

Language of publication	English (English abstracts of non-English publications will be included)	Non-English publications
Date of publication	2006 onwards 2011 onwards for conference abstracts	Publications published prior to 2006

5.1.3 Included and excluded studies

In total, 16 studies were identified; 13 investigated the cost-effectiveness of an active therapy versus BSC in patients that had failed prior systemic treatment and provided a cost/QALY. Of the remaining

three studies, one investigated the cost-effectiveness of an active comparator versus another active comparator and the remaining two did not report a cost/QALY.

Of the 13 studies that reported a cost/QALY with an active comparator versus BSC, three were conducted in the UK from the persepective of the NHS and were therefore considered the most relevant to the decision problem.

- Hoyle et al.[33] reported that compared to BSC, sorafenib treatment resulted in an incremental cost per QALY gained of £75,398, based on an estimated mean gain of 0.27 QALYs per patient .
- The ERG assessment of the cost-effectiveness analysis performed for the everolimus STA submission[4] reported a cost/QALY of £65,231 for everolimus + BSC versus BSC alone (with a PAS scheme applied) compared with the cost/QALY of £51,613 for everolimus + BSC versus BSC alone (with a PAS scheme applied) sumbitted by the manufacturer.
- Thompson-Coon et al.[3] reported a cost/QALY of £102,498 for sorafenib vs BSC.

A summary of all the identified studies is presented in Appendix 2.

5.1.4 Conclusions of the cost effectiveness review

No specific conclusions from the economic review were provided in the MS. The ERG concludes that the identified studies may provide valuable information regarding costs, utilities and model structure, but that they do not negate the necessity of developing a *de novo* model for the current comparison.

5.2 Summary and critique of manufacturer's submitted economic evaluation by the ERG

Table 5.2 presents a summary of the *de novo* economic model developed by the manufacturer.

Table 5.2: Summary of the manufacturer's economic evaluation

	Approach	Source / Justification	Signpost (location in MS)
Model	Markov model with transition probabilities based on PFS and OS curves.		Section 7.2.5 (p.132)
States and events	Three health states are distinguished: progression-free, progressed disease or death.	The model structure and the health states utilised are typical of modelling in metastatic oncology and have been utilised in numerous NICE STAs and MTAs previously.	Section 7.2.2 , 7.2.3, 7.2.4, 7.2.5 (p. 132)
Comparators	Best supportive care	NICE does not recommend any second-line treatment for advanced/mRCC	Section 7.3.1.2 (p. 135)
Natural History	In this study, the approach is the mirror image of the common approach, i.e. the treatment arm (axitinib) is first estimated, and then the treatment effectiveness is used to estimate the BSC arm (i.e. natural history)	Axitinib PFS and OS were based on AXIS trial	Section 7.3.2.1 (p. 135)
Treatment effectiveness	Treatment influences PFS and OS compared to BSC	In cytokine refractory subgroup, BSC PFS and OS were based on indirect treatment comparison using AXIS and TARGET trials. In the sunitinib refractory subgroup, BSC PFS and OS were based in simulated treatment comparison using AXIS and RECORD-1 trials.	Section 7.3.2.2 and 7.3.2.3 (p. 142)
Adverse events	Included in model as 1 time event, they only have impact on costs, not on utility	Based on observed grade 3/4 AEs with occurrence in more than 5% of the population in the AXIS trial for the axitinib arm, and judged by clinical expert to have implication for resource use. For the BSC group, data from the TARGET study was used.	Section 7.5.7 (p. 160)

	Approach	Source / Justification	Signpost (location in MS)
Health related QoL	Health states specific utility values were estimated. Before and after progression utilities were assumed to be treatment independent	Utilities are based on EQ-5D as administered in AXIS trial. Valuation of EQ-5D health states was based on US study. BSC was assumed the same as axitinib based on expert opinion.	Section 7.4 (p. 153)
Resource utilisation and costs	Treatment cost (adjusted for relative dosing intensity observed), health state cost Progression free (accounting for GP visits, CT scan and blood test) and Progressed disease (accounting for GP visits, specialist community nurse and pain medication), costs of death, and costs of adverse events.	Based on UK reference costs, literature and expert opinion	Section 7.5.5 (p. 157) to 7.5.7 (p. 160)
Discount rates	A 3.5% discount rate was used for both costs and effects	According to NICE reference case	Section 7.2.6 (p. 133)
Sub groups	Two subgroups were considered, cytokine refractory patients and sunitinib refractory patients.	Clinical experts have indicated that these two subgroups are distinct groups of patients	Section 7.2.7 (p. 134)
Sensitivity analysis	One-way deterministic sensitivity analysis, scenario analyses and probabilistic sensitivity analysis	Ranges based on observed confidence intervals, expert opinion and assumptions	Section 7.6 (p. 161)

The ERG has assessed the manufacturer's economic evaluation using the Philips et al. checklist for quality assessing decision analytic models.[34] This is shown in Appendix 3 and is used to assist the narrative critique in the following sections.

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.3: Comparison of the MS model with the NICE reference case

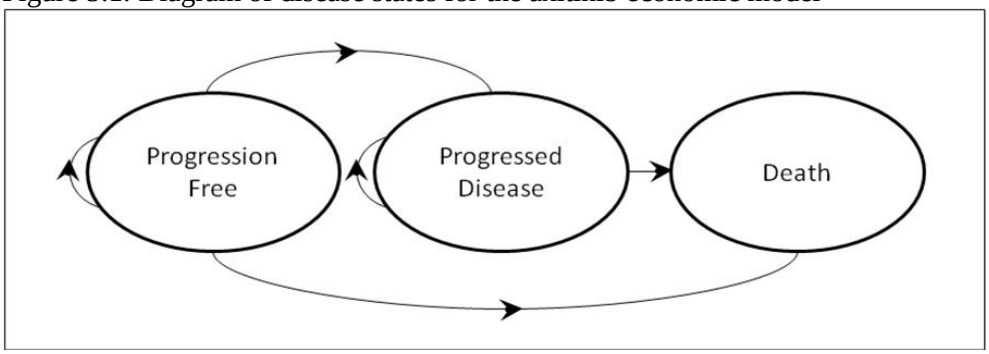
Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether de-novo evaluation meets requirements of NICE reference case
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	Yes	
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	Time horizon 10 years, at that point 3% still alive
Synthesis of evidence on outcomes	Systematic review	Yes	An indirect comparison was done for the cytokine refractory subgroup and a Simulated Treatment Comparison for the sunitinib refractory subgroup
Measure of health effects	QALYs	Yes	
Source of data for measurement of HRQL	Reported directly by patients and/or carers	Partly	EQ-5D administered in axitinib and sorafenib patients, utilities also applied to best supportive care
Source of preference data for valuation of changes in HRQL	Sample of public	Yes	The valuation was based on a US tariff, rather than a UK tariff
Discount rate	Annual rate of 3.5% on costs and health effects	Yes	
Equity weighting	No special weighting	Yes	
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	

5.2.2 Model structure

The cost-effectiveness of axitinib compared to BSC is evaluated using a Markov model. The manufacturer chose the structure of the model based on previously identified models of advanced/mRCC treatment and it was validated by UK clinician expert opinion.[8, 10, 15, 35] It contains the three most relevant health states: progression-free (PF), progressed disease (PD) and death. The PF health state is the period prior to disease progression, whereas the PD state captures the phase post disease progression and prior to death. These health states are those typically utilised in the modelling of metastatic oncology. At any point in time, a patient is assumed to be in one of the states. Patients may move to other states during each four-week model cycle. All patients enter the model in the

progression free health state, having progressed on a previous advanced/mRCC treatment. Patients remain in the progression free health state until they experience disease progression or die. Once patients enter the PD state, they remain there until death. A diagrammatic representation of the model is presented in the Figure 5.1 below.

Figure 5.1: Diagram of disease states for the axitinib economic model



The model uses four week model cycles and a lifetime horizon of 10 years; half-cycle correction is applied to the model. Life years and QALYs gained were generated for the axitinib and BSC arms in order to estimate the incremental cost per QALY gained. The model uses estimates of clinical effectiveness, costs and health related quality of life (HRQoL) estimates to model progression of disease and cost-effectiveness over time. The proportion of patients in each health state at each point in time is calculated directly from parametric survival function equations.

Comment

Since the model chosen by the manufacturer is generally used in the modelling of metastatic oncology and, as mentioned above, is based on previously identified models of advanced/mRCC treatment, the ERG has no comments regarding the model structure.

5.2.3 Population

The manufacturer states that axitinib is indicated for the treatment of adult patients with advanced/mRCC who have received prior systemic treatment. Patients are stratified based on prior treatment regimen, with the majority of patients having received sunitinib or cytokines as their first-line treatment and thus reflecting the licensed indication for axitinib. These two subgroups are examined in separate analyses as cytokine refractory patients are considered by many clinicians to comprise a different subgroup of patients compared with those who are sunitinib refractory.

Comment

The ERG agrees that the population considered is reflective of the actual clinical population and wants to emphasize that approximately 94% of the population of interest will receive sunitinib for first-line treatment and approximately 6% will be treated with cytokines (*see MS page 24*).

5.2.4 Interventions and comparators

There are no therapies currently recommended by NICE for second-line treatment of patients with advanced/mRCC for whom first-line therapy has failed. Therefore, for each of the two populations mentioned in the previous section, the main comparator in the manufacturer submission is best supportive care (BSC) which is in line with the scope and current NICE guidance.

Comment

Both the intervention and the comparator are in line with the final scope.

5.2.5 Perspective, time horizon and discounting

The analysis performed in the manufacturer submission was conducted from an NHS and Personal Social Services perspective in England and Wales using a lifetime horizon of 10 years, with 3.5% per annum discounting applied for cost and QALY benefits. The model allows shorter and longer time durations, and higher and lower discount rates, for sensitivity analysis.

The ERG requested in the clarification letter (Section B – Question 1) that the manufacturer would motivate why a time horizon of 15 years, which is more in line with the real life expectancy, was not chosen as the base case. In their response, the manufacturer stated that:

> *"A 10 year time horizon was chosen as the base case in keeping with previous technology appraisals in mRCC. This is a conservative assumption as the scenario with a time horizon of 15 years which was explored in sensitivity analysis resulted in more favourable ICERs for axitinib."*

For example, the base case (10 years) ICER for the cytokine refractory subgroup (with PAS) is £65,326 and the 15-year ICER is £64,359 (from Table 57 in the MS). Therefore, the choice of a time-horizon of 10 or 15 years has a minimal impact on the model.

Comment

The ERG concludes that the discount rates and perspectives are in line with the NICE reference case. In the MS, it is stated that at 10 years approximately 3% of patients are still alive in the health economic model. Whether this time horizon could thus be considered life time is debatable. However, given the minimal impact of prolonging the time horizon, the ERG considers the 10 year time horizon acceptable.

5.2.6 Treatment effectiveness

As in the manufacturer submission, we first discuss the parametric survival analysis carried out to incorporate the axitinib treatment arm in the economic model for the cytokine refractory and sunitinib refractory populations. Next, the results of the indirect comparison methodology used to model BSC in the cytokine refractory population are discussed. Finally, the results of the two methodologies used to model BSC in the sunitinib refractory population are discussed.

5.2.6.1 Axitinib treatment arm – extrapolation approach

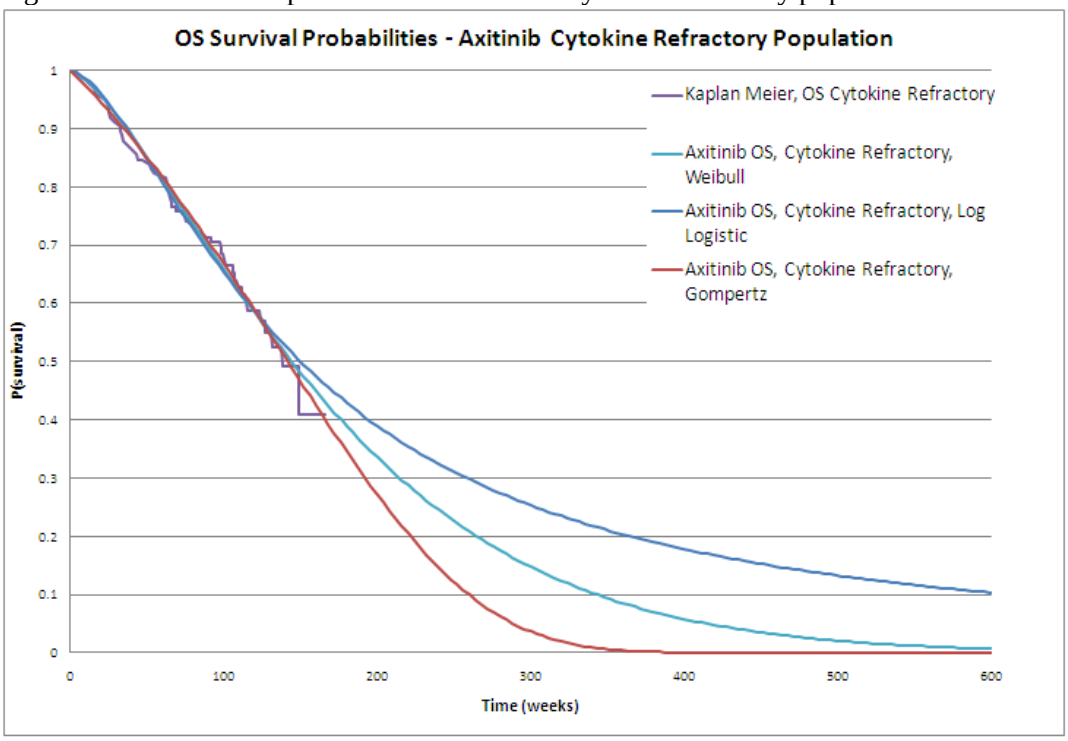
Progression-free survival (PFS) and overall survival (OS) probabilities were incorporated into the model using parametric survival curves. Patient-level data[1] for the cytokine refractory and sunitinib refractory subgroups were analysed for goodness-of-fit using the Exponential, Weibull, Gompertz, Lognormal and Log-logistic distributions. Of these five distributions, the three judged the best fits were included in the model, with the base case representing the most plausible survival estimate. To determine the best model fit, the following criteria were considered, with the most appropriate model identified based on a combination of these: Akaike's information Criteria (AIC) and Bayesian Information Criteria (BIC) statistics; visual inspection and anchoring (i.e. comparison of extrapolation estimates with external data sources). For more details on these criteria we refer to MS page 138.

Prior cytokine - OS

The Weibull model was chosen for the base case, with Log-logistic and Gompertz explored in scenario analyses. As Figure 5.2 shows, the Weibull model provides an intermediate survival estimate between Log-logistic and Gompertz.

[1] Patient level data on PFS and OS were based on the most recent June 2011 and November 1, 2011 data cut-off respectively.

Figure 5.2: OS Survival probabilities – Axitinib cytokine refractory population



Furthermore, the manufacturer states that high-quality anchoring data were available from an axitinib Phase II study in a cytokine refractory population, where the five year survival rate from this study (20.6%, 95% CI 10.9-32.4) corresponded almost exactly to the five year Weibull prediction (20.8%), with the Gompertz and Log-logistic estimates (9.9% and 29.8% respectively), corresponding closely to the upper and lower confidence intervals. Table 5.4 shows the model fit of the survival functions in terms of AIC and BIC statistics.

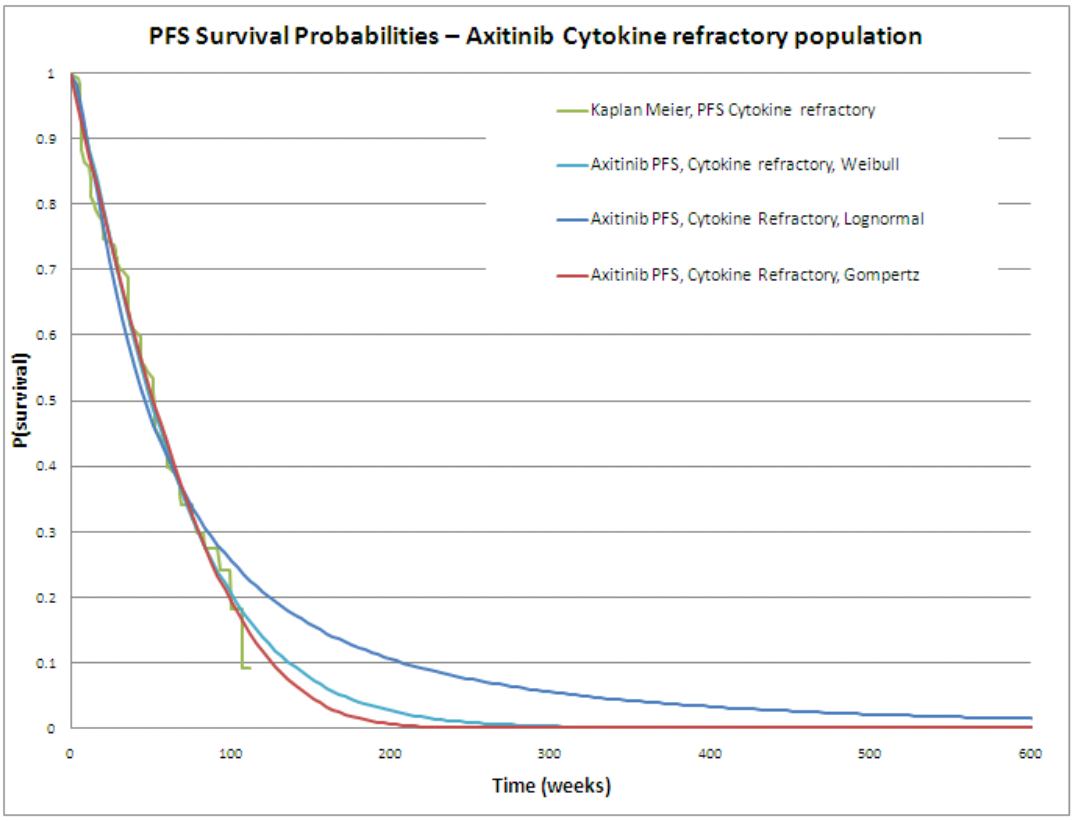
Table 5.4: Model fit for OS in the cytokine-refractory population

Model	Degrees of freedom	AIC	BIC
Weibull	2	250.1823	255.8548
Gompertz	2	251.2509	256.9235
Log-logistic	2	250.7399	256.4124

Prior cytokine - PFS

The Weibull curve was again chosen as the base case, with Gompertz and Lognormal presented as scenarios (see Figure 5.3).

Figure 5.3: PFS Survival probabilities – Axitinib cytokine refractory population



For the PFS extrapolation, less variation between the different models was seen than for OS. The MS explains that this is due to the higher proportion of patients having reached the PFS endpoint during the follow-up period than in the OS data. Table 5.5 shows the model fit of the survival functions in terms of AIC and BIC statistics.

Table 5.5: Model fit for PFS in the cytokine-refractory population

Model	Degrees of freedom	AIC	BIC
Weibull	2	293.5021	299.1747
Gompertz	2	294.2111	299.8837
Lognormal	2	293.7307	299.4033

In this case, the MS states that the lognormal model predicted a higher proportion of non-progressed patients at 10 years, as can be seen in Figure 5.3. This was felt to be clinically implausible by the experts since, according to the response to the clarification letter (Section B – Question 11), "*a small proportion of mRCC patients may survive for long periods due to the heterogeneous nature of the disease, meaning that a distribution with a tail at 10 years may be plausible for OS"*. However, this is not the case for PFS as "*it would not be expected for patients to continue on treatment with axitinib for a time period as long as 10 years*".

Nevertheless, the manufacturer decided to explore the lognormal model as additional scenario.

Prior sunitinib - OS

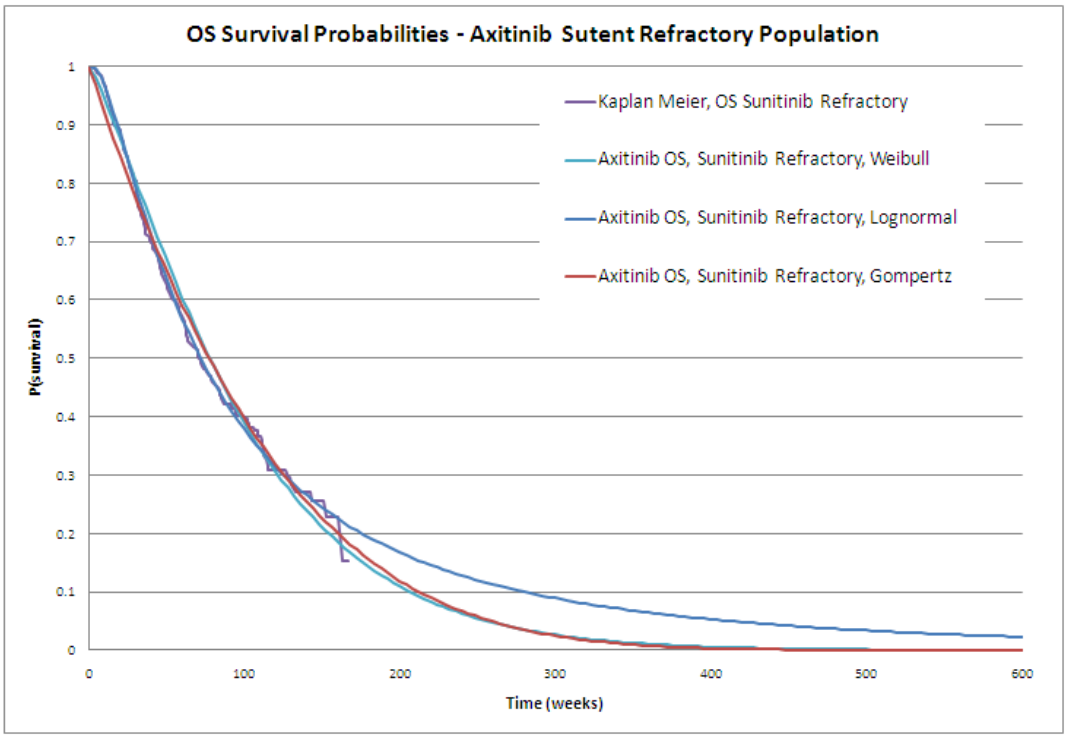
The manufacturer states that in contrast to the cytokine refractory population, where a larger proportion of patients remained alive at the end of the trial follow up period, the sunitinib refractory dataset was more complete and allowed for more accurate OS extrapolation. In this case, the Lognormal provided the most accurate fit to the data and the lowest AIC/BIC values, as can be seen in Table 5.6. Moreover, it was considered a more clinically plausible survival estimate by the experts consulted, as the tail of the curve is more consistent with daily practice, where a low proportion of treated patients can be expected to live a longer period. Therefore, the Lognormal model was chosen as the base case, with Weibull and Gompertz examined in scenario analyses. In this case, as mentioned in the response to the clarification letter (Section B – Question 12), the log-logistic and lognormal distributions showed very similar fits. The Gompertz model was chosen in order to illustrate the impact of a more conservative survival estimate in the model.

Table 5.6: Model shapes for OS in the sunitinib-refractory population

Model	Degrees of freedom	AIC	BIC
Weibull	2	506.633	513.1687
Gompertz	2	512.2575	518.7933
Lognormal	2	496.1517	502.6874

The three survival curves can be seen in Figure 5.4. The MS discusses how the sunitinib-refractory Kaplan Meier curve seems to show a non-monotonic hazard, with the curve appearing more concave in the middle portion. The Gompertz and Weibull curve appear to over-predict survival in the middle part of the curve, whereas the Lognormal tracks the curve better for the entire period.

Figure 5.4: OS Survival probabilities – Axitinib sunitinib refractory population



Prior sunitinib - PFS

The Weibull was chosen as the base case, with Lognormal and Gompertz curves included in the model as scenarios. The three models provided similar fits, as can be observed in Figure 5.5. According to the MS this is due to the fact that the survival data were over 90% complete at the cut-off date.

Figure 5.5: PFS survival probabilities – Axitinib sunitinib refractory population

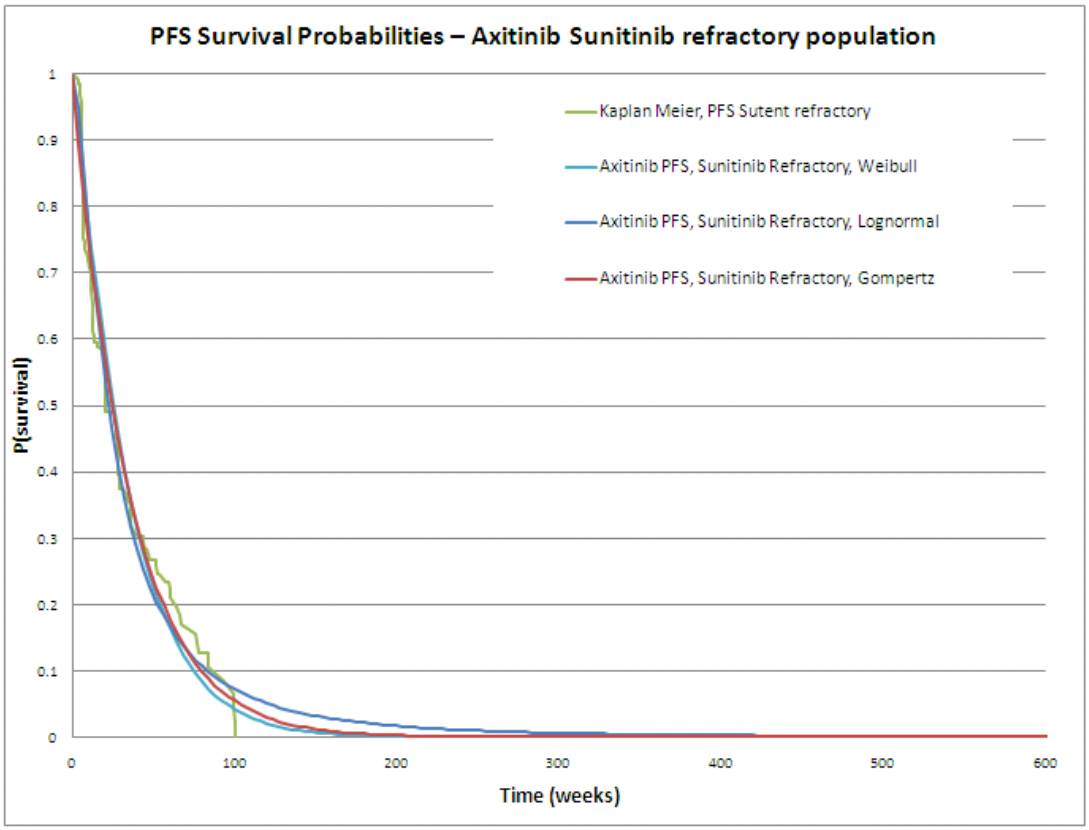


Table 5.7 shows that the Lognormal curve had the best fit in terms of AIC and BIC. However, the survival estimate at the tail end of the curve was considered clinically implausible. Therefore, the Weibull model, which produced an intermediate PFS estimate between Lognormal and Gompertz, was chosen as base case.

Table 5.7: Model shapes for PFS in the sunitinib-refractory population

Model	Degrees of freedom	AIC	BIC
Weibull	2	496.7759	503.3116
Gompertz	2	498.9336	505.4693
Lognormal	2	475.3779	481.9136

5.2.6.2 BSC comparison – cytokine refractory population

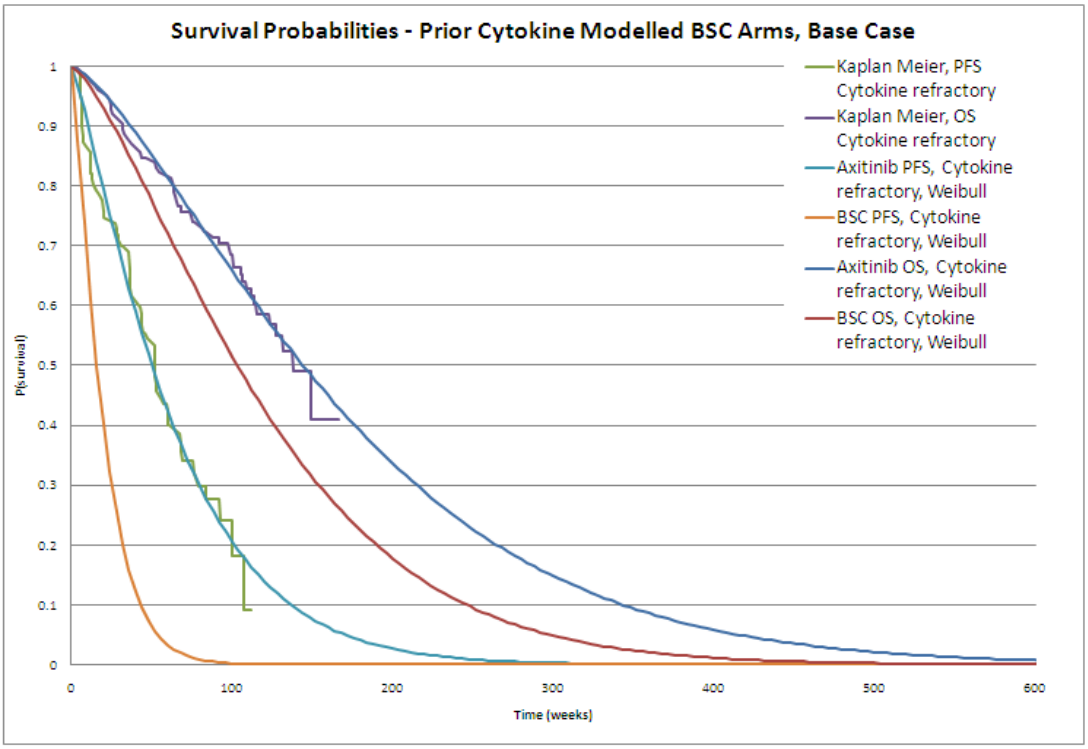
To model the BSC arm for the cytokine refractory population, the manufacturer chose an indirect comparison between axitinib and BSC via the TARGET crossover-censored hazard ratio, as it was identified as the most valid methodological approach (see Section 4.2.3). The results of the indirect comparison in terms of hazard ratios can be seen in Table 5.8 and the modelled BSC curves are shown in Figure 5.6.

Table 5.8: Axitinib-BSC cytokine refractory hazard ratios used in the economic model

	HR (95% CI): Axitinib vs BSC)
PFS	0.251 (0.165-0.379)
OS	0.63 (0.41-0.99)

Abbreviations: BSC, best supportive care; CI, confidence interval; OS, overall survival; PFS, progression-free survival.

Figure 5.6: Survival probabilities – Prior cytokine modelled BSC arms, base case



5.2.6.3 BSC comparison – sunitinib refractory population

The MS identifies three key limitations of the evidence network for making a comparison between axitinib and BSC in the sunitinib refractory group. These are the following: the uncertainty in the incremental OS measurement from the AXIS study, the confounding of OS data in the TARGET study due to cross-over and the lack of evidence comparing sorafenib to BSC in a prior sunitinib population. Given these limitations, both the RENCOMP and STC methodologies, described in Section 4.2.6, were examined in the modelling approach. The MS states that the RENCOMP analysis replaces the gap in the evidence network since it provides clinical data in a prior sunitinib patient population who received second-line sorafenib or BSC. This analysis corrects only two of the shortcomings mentioned above. The STC, however, overcomes all these three limitations since it allows a direct link to be made between the AXIS axitinib arm and the RECORD-1 BSC arm, removing the requirement to correct for confounding in the AXIS OS relationship. Therefore, the manufacturer chose the STC as the base case approach for the sunitinib refractory population, and explored the RENCOMP in a scenario analysis.

Simulated treatment comparison

Two data sources, denoted by RECORD-1 ITT placebo population and RECORD-1 prior sunitinib everolimus, were examined in the STC.

The results of the calibrated PFS for ITT RECORD-1 placebo data are shown in Figure 4.2 and Figure 4.3 for the Lognormal and Weibull distributions, respectively. The calibrated PFS for RECORD-1 prior sunitinib everolimus data is shown in Figure 4.4 for the Weibull distribution. The prior sunitinib PFS hazard ratio from the RECORD-1 study (HR=0.34; 95% CI: 0.23-0.51) was applied to the everolimus STC curve to generate a modelled AXIS-like, prior sunitinib PFS curve. As the Lognormal model does not support the application of a hazard ratio, the Weibull was the only option explored in the model. A summary of predicted STC survival times for PFS is presented in Table 4.12.

The results of the calibrated OS for ITT RECORD-1 placebo data are shown in Figure 4.5 and Figure 4.6 for the Lognormal and Weibull distributions, respectively. The calibrated OS for RECORD-1 prior sunitinib everolimus data are shown in Figure 4.7 for the Weibull distribution. To create a modelled placebo arm for the everolimus prior sunitinib population, the RPSFT-adjusted OS hazard ratio from the RECORD-1 study (equal to 0.53) was applied to the AXIS-like everolimus curve to generate a modelled AXIS-like, sunitinib refractory placebo curve. The RPSFT-adjusted hazard ratio was chosen as it was validated by the NICE ERG during the everolimus appraisal and was used to derive the final OS estimate included in the everolimus economic model.[4] As the Lognormal model does not support the application of a hazard ratio, the Weibull was the only option explored in the model. A summary of predicted STC survival times for OS is presented in Table 4.14.

RENCOMP Indirect comparison

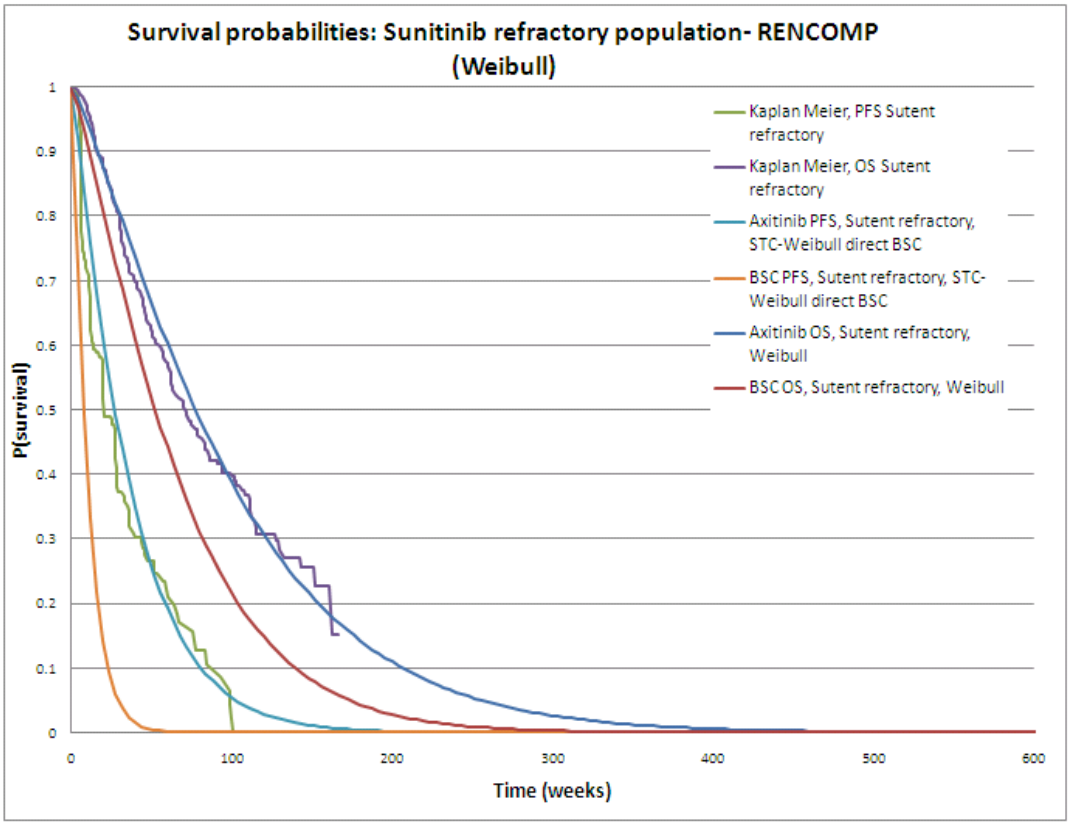
The manufacturer carried out an indirect comparison by incorporating the RENCOMP hazard ratio into a meta-analysis, using the sunitinib refractory OS hazard ratio from the AXIS study (0.997, 95% CI 0.782,1.27) to generate an axitinib-BSC OS hazard ratio. The calculated hazard ratio, with a 95% confidence interval, can be seen in Table 5.9.

Table 5.9: Axitinib-BSC sunitinib refractory (via RENCOMP) OS hazard ratio.

	OS HR (95% CI)
Axi-BSC RENCOMP HR	0.619 (0.384-0.997)

The above hazard ratio is incorporated into the economic model through parametric survival curves for best supportive care generated by applying the OS hazard ratio to the axitinib parametric survival function, as described in the cytokine refractory section. Figure 5.7 illustrates the modelled BSC survival function using the RENCOMP hazard ratio when applied to the Weibull model.

Figure 5.7: Survival probabilities – RENCOMP (Weibull)



The MS argues that despite the better fit provided by the Lognormal model, the Weibull was used as the base case for the RENCOMP data, as accelerated failure time models like the Lognormal and Log-logistic assume a constant proportional hazard and do not allow for the application of a hazard ratio into the functional form. Nevertheless, the application of the hazard ratio to the Log-logistic model was explored in scenario analysis using the functional approach detailed in appendix 19 of the MS.

5.2.6.4 Summary of survival-probability-related parameters used in the economic model

Table 5.10: Survival-probability related parameters of the model

Variable	**Value**	**Measurement of uncertainty and distribution**
Hazard Ratios	HR	95% CI (lognormal)
Prior cytokine		
PFS, Axi vs. BSC via TARGET	0.251	0.165 - 0.379
OS, Axi vs. BSC via TARGET crossover-censored	0.63	0.41 - 0.99
Prior sunitinib		
OS, Axi vs BSC via RENCOMP model	0.619	0.384 - 0.997
Survival Function Parameters	Value	Covariance Matrix
Axitinib cytokine refractory – PFS		
Weibull – Axitinib Parameter 1 = λ	[redacted]	**ln(λ)** **ln(γ)** **ln(λ)** [redacted]

Variable	Value	Measurement of uncertainty and distribution		
Parameter 2 = γ		**ln(γ)**	[redacted]	[redacted]
Lognormal – Axitinib Parameter 1= mean μ Parameter 2= ln(σ)	[redacted]		**Const**	**ln (σ)**
		Const	[redacted]	
		ln (σ)	[redacted]	[redacted]
Gompertz – Axitinib Parameter 1 = const Parameter 2 = γ	[redacted]		**Const**	**γ**
		Const	[redacted]	
		γ	[redacted]	[redacted]
Axitinib cytokine refractory – OS				
Weibull Parameter 1 = λ Parameter 2 = γ	[redacted]		**ln(λ)**	**ln(γ)**
		ln(λ)	[redacted]	
		ln(γ)	[redacted]	[redacted]
Log-logistic Parameter 1= λ Parameter 2= ln(γ**)**	[redacted]		**Const**	**ln(γ)**
		Const	[redacted]	
		ln(γ)	[redacted]	[redacted]
Gompertz Parameter 1 = const Parameter 2 = γ	[redacted]		**Const**	**γ**
		Const	[redacted]	
		γ	[redacted]	[redacted]
Axitinib sunitinib refractory – PFS				
Weibull Parameter 1 = λ Parameter 2 = γ	[redacted]		**ln(λ)**	**ln(γ)**
		ln(λ)	[redacted]	
		ln(γ)	[redacted]	[redacted]
Lognormal Parameter 1 = mean μ Parameter 2 = S= ln(σ)	[redacted]		**Const**	**ln (σ)**
		Const	[redacted]	
		ln (σ)	[redacted]	[redacted]
Gompertz Parameter 1 = const Parameter 2 = γ	[redacted]		**Const**	**γ**
		Const	[redacted]	
		γ	[redacted]	[redacted]
STC Adjustment Factors – PFS				
Weibull via sunitinib refractory via BSC	[redacted]			
Lognormal via sunitinib refractory via BSC	[redacted]			
Weibull via sunitinib refractory via everolimus	[redacted]			
Axitinib sunitinib refractory – OS				
Weibull	[redacted]		**ln(λ)**	**ln(γ)**

Variable	Value	Measurement of uncertainty and distribution		
Parameter 1 = λ Parameter 2 = γ	[redacted]	**ln(λ)**	[redacted]	
		ln(γ)	[redacted]	[redacted]
Lognormal Parameter 1 = mean μ Parameter 2= S= ln(σ)	[redacted]		**Const**	**ln (σ)**
		Const	[redacted]	
		ln (σ)	[redacted]	[redacted]
Gompertz Parameter 1 = const Parameter 2 = γ	[redacted]		**Const**	**γ**
		Const	[redacted]	
		γ	[redacted]	[redacted]
STC Adjustment Factors				
Weibull via sunitinib refractory via BSC	[redacted]			
Lognormal via sunitinib refractory via BSC	[redacted]			
Weibull via sunitinib refractory via everolimus	[redacted]			

Abbreviations: AE, adverse event; BSC, best supportive care; CI, confidence interval; CT, computed tomography; GP, general practitioner; HR, hazard ratio; OS, overall survival; PFS, progression-free survival; STC, simulated treatment comparison.

Comment

In general, the ERG has few comments on the way the PFS and OS were estimated. The manufacturer followed the approach outlined in the DSU report on the extrapolation of survival curves,[29] and in general, reasonable choices were made. The ERG does however have some minor comments.

For example, the Gamma distribution was not included as part of the set of distributions used for analysis. In their response to the clarification letter (Section B – Question 3) the manufacturer provided the following two arguments for this:

- The difficulty of implementing the Gamma distribution in the Excel 2007 model framework (it would have required complex VBA coding to incorporate a probabilistic analysis).
- The fact that the Gamma distribution is an accelerated time failure (AFT) model.

The ERG agrees with the first reason but not with the second since the manufacturer used log-logistic and lognormal models which are also AFT models.

In addition, the ERG considers the procedure to select the base case model unclear. This procedure is based on the combination of three different criteria (AIC/BIC, visual inspection and anchoring) and according to the manufacturer "*no one factor was viewed as dominant to another when choosing extrapolation options. However, as the long-term survival trends are a key driver of the model results, the plausibility of long-term survival estimates was an important factor and in many cases overrode other factors*". This seems to be contradictory. Moreover, from the descriptions in MS Section 7.3.2 it appears that expert opinion was always dominant in the model selection. This seems to be confirmed by the manufacturer in the response to the clarification letter (Section B – Question 5): "*The plausibility of survival estimates were determined based on clinical opinion, knowledge of the natural history of renal cell carcinoma, examples from previous NICE appraisals and other HTA appraisals for late-stage metastatic solid tumours, and clinical and product knowledge of sunitinib and axitinib*".

Moreover, the manufacturer states in their response that "*when several distributions demonstrated similar fits for all criteria, "high, medium and low" survival estimates were chosen for the economic model, to allow for examination of the impact of different survival assumptions on the model outcome. Additionally, in the cases when survival in the other treatment arm was modelled through the use of a hazard ratio (as was the case for the cytokine refractory subgroup), only the best fitting of the two accelerated failure time models was retained in the scenario analysis*", i.e., "*the list of 3 best fitting distributions has to include at least 2 proportional hazard models"*. These reasons are clear but never mentioned in the MS.

Overall, for three of the four curves that were fitted, the final choice coincided with the curve showing the best statistical fit. Only for the PFS curve in the sunitinib refractory group led expert opinion to a base case distribution that did not have the best statistical fit, but was clinically more plausible. In conclusion, the ERG accepts the choices made for base case distributions and distributions considered in the scenario analyses.

In the clarification letter (Section B – Question 2), the ERG asked for the rationale for assuming that patients who withdraw from treatment continue to follow the PFS and OS curves for axitinib, rather than following the PFS and OS curves of the BSC group after withdrawal.

In their response, the manufacturer states that, in line with the AXIS study, patients who discontinued the treatment before progression due to adverse events were still followed up in the trial and included in the estimation of the PFS and OS curves for the axitinib arm. For that reason, it would not be appropriate to assume that axitinib patients who discontinue treatment would have similar PFS and OS to BSC patients.

The ERG does agree that this is a valid approach, as long as patients who discontinued treatment were followed up not only regarding survival but also progression. However, from the clinical study report, we were not able to find conformation that patients withdrawing from treatment due to adverse events were still followed for progression. From section 5.4.1 of the clinical study report it appears that these patients were taken into account in the PFS curve only until the date of treatment withdrawal:

> *"PFS data were censored on the date of the last tumor assessment (on-study) documenting absence of PD for patients*
> - *Who had at least 1, on-study disease assessment and discontinued treatment without documented disease progression and without death on-study;"*

If this is indeed the case, the estimate of the QALYs in the axitinib group is affected, as these patients will progress earlier once off treatment. However, the number of life years will not be affected, as death was recorded also for patients who withdrew prematurely. Additionally, the costs in the axitinib group is affected as earlier transition to the progressed health state means higher overall costs (compared to the costs currently in the model, i.e. in which treatment costs of these patients is set to zero). Taken together, the ERG expects that if it were possible to correct this issue, the ICER would become slightly larger, though the impact is limited by the relatively small group of patients withdrawing from treatment prematurely.

More theoretically, the ERG wants to point out that this approach of the manufacturer could lead to a model that is a good reflection of the trial data (assuming patients are followed for progression after treatment withdrawal), but that does not allow for assessing the impact of a smaller or larger percentage of treatment discontinuation. In the current model, such scenario analysis would only

impact the costs, as more treatment withdrawal leads to lower treatment costs while still achieving the same effects, thus leading to lower ICERs.

Thus, the ERG feels that ideally, a model would make the impact of treatment withdrawal explicit, not only for the costs, but also for the time-to-event. However, we recognize that due to lack of data, this will often not be possible.

Finally, as mentioned in Section 4.2.6, an error was found in the calculation of the PFS hazard ratio for the cytokine refractory population. This has a minimal impact on the results of the analyses shown in the MS, decreasing the ICER slightly. The results obtained with the correct value of the PFS hazard ratio for the cytokine refractory population can be seen in Section 5.3.

5.2.7 Health related quality of life

Base case – AXIS study

The manufacturer collected utility data in the AXIS trial using the EuroQoL-5D (EQ-5D) instrument, completed by the study patients at Day one, every four weeks afterwards, at the end of study treatment or withdrawal and at follow up Day 28 (28 days after the last dose of active treatment). The quality of life analysis was based on the ITT population (the full analysis set).

The baseline mean (SD) EQ-5D score (Day one of Cycle one) for the axitinib arm was 0.732 (0.01). The mean on-treatment utility for axitinib was calculated by averaging the EQ-5D index value at each time point, weighted by the number of patients still on treatment at that time point, giving a mean (SD) utility of 0.692 (0.275). Finally the progressive disease utility was based on the average end-of-treatment utility, giving a mean (SD) utility of 0.610 (0.316). Note that all these estimates were based on the axitinib and sorafenib groups together.

The MS highlights the absence of relevant literature on utility values for treatment with BSC. Therefore, the manufacturer made the assumption that BSC patients would experience the same utility as patients receiving active treatment with axitinib. This assumption, tested and confirmed by the main clinical advisor for the economic model, was argued as follows: while patients with axitinib may expect to experience some reduction in health-related quality of life related to the treatment, they will also receive HRQoL benefit in terms of symptomatic control and disease stabilization.

Scenario analysis – previous NICE utility estimates

The manufacturer also carried out a scenario analysis with the utility figures used in several previous NICE appraisals to model second-line mRCC. As these utility estimates are based on consensus between UK experts, the NICE appraisal committee and ERG groups from several appraisals, they allow for "like versus like" comparability between axitinib and other previous advanced/mRCC appraisal.

Adverse events

The MS states that the HRQL estimates included in the AXIS trial reflect the adverse event profile associated with axitinib. Therefore, the utility estimates included in the economic model are expected to reflect the adverse event profile of the treatment. Thus, no specific utilities were included to model adverse events.

Quality-of-life data used in cost-effectiveness analysis

A summary of the QoL values used in the economic analysis is presented in Table 5.11.

Table 5.11: Summary of quality of life values for cost-effectiveness analysis

	State	Utility value, mean (SD/SE)	Reference to section in submission
Base case	Progression Free	0.692 (0.275)	AXIS - weighted mean on-treatment utility for axitinib patients (7.4.3.1)
	Progressed	0.610 (0.316)	AXIS – mean utility at treatment discontinuation (7.4.3.1)
Scenario analysis	Progression Free	0.758 (0.03)	Previously utilised utility estimates from NICE 2nd-line advanced/mRCC appraisals (7.4.3.2)
	Progressed	0.683 (0.04)	

In the clarification letter (Section B – Question 15) the manufacturer was asked to provide the utility estimates for the progression free and progressive disease health state for the prior sunitinib and the prior cytokine subgroups separately. These, and the p-value from the hypotheses that these utilities are the same in each subgroup, can be seen in Table 5.12.

Table 5.12: Subgroup specific utilities

	Overall	Sunitinib refractory	Cytokine refractory	p-value
PF	0.692	[redacted]	[redacted]	[redacted]
PD	0.61	[redacted]	[redacted]	[redacted]

Comment

As mentioned above, the MS assumes no difference in utilities between the two treatment arms. However, the study by Swinburn et al. (2010)[36] shows that grade III diarrhoea, grade III fatigue and grade III hypertension has a significant impact on the utility score for stable disease from 0.795 to 0.534, 0.591 and 0.642 respectively [TTO with UK members of the general public]. Moreover, the study by Zbrozek et al. (2010)[37] also shows the influence of serious toxicity on the utility score. Therefore, the manufacturer was asked in the clarification letter (Section B – Question 17) to provide further justification of this assumption. The manufacturer states that, although the study by Swinburn et al, 2010[36] certainly "*found considerable disutilities for selected adverse events when compared to likely utility of patients on stable disease receiving first line mRCC treatment*", the impact of adverse events on the mean utility value per patient on treatment would not be high since only a small proportion of patients receiving mRCC treatment will experience AEs. Moreover, this impact "*is expected to be applicable only for the period where the AE is experienced. As most grade 3-4 AEs will be actively managed and resolved either through urgent clinical intervention or dose reduction/interruption it is expected that the duration of the utility decrement due to the AE will be short*". In addition, the AEs utility values estimated in Swinburn et al., 2010[36] are compared to the utility of patients receiving first line treatment of mRCC. Therefore, the baseline utility is expected to be lower in the second line treatment. Finally, the utility data reported in AXIS, shows that the mean utility for patients in the axitinib was 0.692 compared to a mean baseline utility of 0.732, which reflects a utility decrement (on average) of 0.04 for the axitinib arm. On the other hand, patients on the BSC arm would experience a reduction in HRQoL due to disease symptoms and actively progressing uncontrolled disease. As the manufacturer states, "*this negative impact might be less significant when compared to that of some serious adverse events but it might have the same overall*

impact on the mean utility per patient as it is likely to affect more patients. This is further supported by QoL evidence comparing placebo with active treatment in 2nd line mRCC, which suggests that QoL is similar to placebo despite the presence of AEs".

The ERG does agree that this is a valid assumption and that patients receiving best supportive care may experience the same utility as patients receiving treatment with axitinib. Nevertheless, the ERG explored the impact of having different utilities for the different treatment arms in additional scenario analyses. These results are presented in Section 5.3.

For the progressive disease health state, the utility was based on the end-of-treatment EQ-5D questionnaire, i.e. the utility reflects the quality of life at the time the patient enters this health state. In the model, the utility then remains constant whereas in reality, it would be reasonable to assume that utility declines as patients near their end of life. So in reality, the average utility in the progressive disease phase may be expected to be lower than the utility at the start of that phase, as is currently assumed. This would increase the ICER slightly when axitinib patients stay longer in the progressive disease phase than BSC patients (as is the case for sunitinib-refractory patients, see table 5.16). If however both patient groups spent the same time in the progressive disease phase, a lower utility has no impact on the ICER (as is the case for cytokine refractory patients, see table 5.16).

In the MS, no information was provided on the approach to valuation of the EQ-5D descriptive health states. The clinical study report indicated that health states were valued using a US tariff developed by Shaw et al. [38]. The question arises to what extend the US valuation is similar to UK valuation. Using a dataset in which a TTO was performed on a set of 42 EQ-5D health states both by US and UK respondents, 2 separate comparisons were made, using different statistical methodologies.[39, 40] Both studies concluded that the US valuation is consistently higher than the UK valuation, and this is more so for worst health states. This would imply that the utilities used in the current appraisal are slightly too high (PF 0.692, PD 0.61). However, without patient level data it is difficult to say how much the utility estimates should be lower. So, given that the ERG does not have such patient level data and since these values are already clearly lower than the values used in previous NICE appraisals we have not defined alternative values for an ERG base case. However, we did explore the impact of lower utilities in both health states in an additional scenario analysis which is presented in section 5.3.

The ERG found errors in the utilities considered for the base case scenario. The mean PF and PD utilities reported (0.692 and 0.610, respectively) are in fact *sample* means. Therefore, they should be accompanied with a standard error (SE) instead of a SD. With the data provided in the response to the clarification letter (Section A - Question 6), the ERG was not able to reproduce the PF utility, and following the method described in the MS, a higher utility was found, i.e. 0.732. For this higher utility, we found an SE of 0.0035. We have applied this SE instead of the SD of 0.275 in the sensitivity analyses. For PD, the SD reported in Table 5.11 of 0.316 should be replaced by an SE of 0.0175. The use of the SE instead of the SD has a large impact on the uncertainty associated to the utilities and it is reflected in the PSA. This is especially significant for the sunitinib refractory population as can be seen in Section 5.3.

Finally, the ERG explored the impact of using the subgroup specific utilities presented in Table 5.12. The ICERs obtained using these utilities are reported in Section 5.3 in Table 5.25.

5.2.8 Resources and costs

5.2.8.1 Costs of treatment

The manufacturer modelled the cost of axitinib based on the recommended dosing schedule for the product (5mg BD). Since axitinib is orally administered, no administration costs were included in the model. Treatment costs are associated to the PF state only (in keeping with the AXIS trial and recommended UK clinical practice for TKI treatment in advanced/mRCC). Moreover, the model reflects the effective price of axitinib to the NHS without the proposed patient access scheme (PAS). The MS adjusted axitinib cost for the relative dosing intensity (RDI) observed in the AXIS trial (102%) and further included in a probabilistic sensitivity analysis according to the observed standard deviation (35.2%). Additionally, the manufacturer performed a scenario analysis to explore the impact of a lower dosing intensity, since expert opinion and observed clinical practice indicates that real-world dosing intensities are typically lower than those observed in clinical trials. Thus, an RDI of 80% was explored in scenario analysis (in keeping with clinician expert opinion and previous NICE appraisals in second-line advanced/mRCC (everolimus)). Finally, the manufacturer assumed no drug costs for best supportive care. Table 5.13 presents the resulting treatment costs for axitinib.

Table 5.13: Unit costs associated with the technology in the economic model

Items	Intervention (confidence interval)	Justification
Technology cost	[redacted] £3,517/cycle (28 days)	[redacted] Effective cost to NHS (without PAS)
Dosing intensity (base case)	102.0% (SD 35.2%)	Observed dosing intensity in AXIS study
Dosing intensity (scenario analysis)	80%	Intended to explore the impact of lower dosing intensity in real-world clinical practice; consistent with clinical opinion and previous NICE appraisals
Administration costs	n/a	Therapy administered orally with no associated costs for administration

In the clarification letter (Section B – Question 19) the manufacturer was asked to provide the dose intensity for the patients who were sunitinib refractory and cytokine refractory in the AXIS trial separately and to report ICERs with the sub-group specific dose intensities. These are given below. The mean dose intensity for the sunitinib refractory and cytokine refractory patients receiving axitinib in AXIS was [redacted]. The ICERs with the sub-group specific dose intensity for sunitinib refractory patients were £40,639 with PAS and [redacted] without PAS. For cytokine refractory patients the ICER with the sub-group specific dose intensity was £66,955 with PAS and [redacted] without PAS. In addition, the manufacturer was asked to justify the use of the overall dose intensity instead of sub-group specific for the base case. To this latter question no answer was provided.

5.2.8.2 Health-state costs

The estimates of routine medical monitoring for the stable and progressed disease states included in the MS were based on those considered in the PenTAG economic model[3] and the everolimus STA.[4] Since patients are assumed to receive equivalent management independently of the treatment received, health-state costs were applied equally to the axitinib and BSC arms. In the MS all costs are

updated to current values, or inflated using the PSSRU Health Care Inflation Index for Hospital and Community Health Services where recent references were not available.

For the PF health state, the manufacturer considered costs for patient monitoring (one GP visit per cycle), tumour scans (one scan per three cycles), and blood tests (one test per cycle). In particular, for patient monitoring, the manufacturer assumed that patients would receive ongoing management and drug dispensation by GP, as in the everolimus appraisal. Nevertheless, the impact of assuming management by oncologist rather than GP was explored in a scenario analysis. For the PD state, the manufacturer included routine medical management costs for progressive disease for one clinical consultation per month, 1.5 specialist palliative care community nurse visits per month, and pain medications. All these costs are in keeping with the NICE MTA and everolimus submission. In addition, a cost of death of £3,923 was also included in the MS (see MS Table 40), using the reference from Coyle et al. (1999) inflated to 2011 values.[41] Health-state costs included in the manufacturer model can be seen in Table 5.14.

Table 5.14: List of health states and associated costs in the economic model

Health states	Items	Mean frequency or duration	Unit cost (£)	Cost per cycle (£)
Progression free – Base case	GP visit[a]	1 visit per cycle	£53.00/visit	£53.00
	CT scan[b]	1 scan per 3 cycles	£160.00/scan	£53.33
	Blood test[c]	1 test per cycle	£3.36/test	£3.36
Total cost per cycle – Progression free state				£109.69
Progressed disease - Base Case	GP visit[d]	1 visit per cycle	£53.00/visit	£53.00
	Specialist community nurse[d]	3 visits / 8 weeks	£84.00	£126.00
	Pain medication[e]	28 vials per cycle	£5.00/dose	£140.00
Total cost per cycle – Progressed disease state				£319.00
Progression free – Scenario analysis assuming oncologist visits	Oncologist Visit[f]	1 visit per cycle	£120/visit	£120.00
	CT scan	-		As above
	Blood test			As above
Total cost per cycle – Progression free State (Scenario analysis)				£176.69
Progressed disease – Scenario analysis assuming oncologist visits	Oncologist Visit[f]	1 visit per cycle	£120/visit	£120.00
	Specialist community nurse[e]	-		As above
	Pain medication[e]			As above
Total cost per cycle – Progressed disease state (Scenario analysis)				£386.00

Sources: GP visits: Unit Costs of Health and Social Care 2011 (2011), Curtis L[42]

[b]Code RA14Z Computerised Tomography Scan, more than three areas

[c]Code DAP823 Haematology [Excluding Anti-Coagulant Services]

Code 202AF- Band 2 Palliative/respite care: adult face-to-face NHS Trust and PCT combined Reference Costs 2007-08

[e]BNF section 4.7.2 Opioid analgesics (morphine sulphate 1 mg/mL, net price 50-mL vial = £5.00 http://www.medicinescomplete.com/mc/bnf/current/3502.htm#_3502)
[f]Medical Oncology Code 370 for the "National Schedule of Reference Costs Year : 2010-11 - NHS Trusts and PCTs combined Consultant Led: First Attendance Non-Admitted Face to Face"[43]

*In all instances in this table, "Cycle" refers to one 28-day model cycle

In the clarification letter (Section B – Question 21) the ERG asked the manufacturer to describe in detail the costs of death and how the estimate for these costs (before inflating to 2011 costs) was derived from Coyle et al (1999).[41] According to the manufacturer, the mean cost of palliative care in the community, in hospitals and in hospices in the UK was estimated using patient information collected from eight 'district health authorities' located in England and Wales in 1994. These data were available for 231 hospice patients and 95 hospital patients who were recruited into the study within three days of admission to the inpatient setting or within three days of transfer to palliative care. Cost data were collected at baseline, after one week and at monthly intervals for two months. After the third follow up there were no inpatients remaining in the study. Length of stay, procedures received, number and type of tests conducted and other treatments delivered were recorded for each patient in order to estimate the cost of inpatient palliative care. Resource use included bed-day costs, surgical procedures, chemotherapy, radiotherapy, blood transfusions, nasogastric tube insertion, blood tests, ECG, X-rays, ultrasound scans, bone scans, CT and MRI scans. Based on these data, the mean cost of palliative care was estimated to be £2,285.88 (SD £2,096.80) in the hospital and £3,049.91 (SD: £1,791.70) in the hospice setting, respectively. The cost of death in the manufacturer's model was estimated using the weighted average cost (weighted by patient numbers in the study) of inpatient palliative care delivered in either the hospital or the hospice setting was calculated and inflated from 1999 to 2011 using the PSS Pay & prices inflator. Moreover, the manufacturers' model assumed that a cost of death was assigned to the progressed disease state only, and that this cost included hospitalisation costs for palliative care costs as described above.

5.2.8.3 Adverse Events Costs

The manufacturer states in their submission that treatment discontinuation may occur also due to adverse events. This was incorporated to the model by applying a per-cycle rate of adverse event related discontinuation, as explained in the MS section 7.5.5. The costs of adverse events for the axitinib and BSC arms were associated with the PF health state only, and added to the costs of ongoing resource use for this health state. Moreover, it was assumed that AEs were resolved within one cycle. For the axitinib arm, AEs were taken from the ITT population of the AXIS trial and included diarrhoea and hypertension.

For the BSC treatment arm, the MS stated that the AE profiles from the BSC treatment arms of the TARGET trial and RECORD-1 trial were pooled to determine an estimated AE profile for BSC. However, in response to the clarification letter (Section B – Question 24) the manufacturer corrected this statement and explained that they examined two potential sources to inform the AEs for best supportive care. In the RECORD-1 study 5.1% of patients receiving best supportive care experienced anaemia and 0% hypertension, while in TARGET study 2% of patients receiving best supportive care had hypertension and 0% in anaemia. The TARGET study was used in the base case. The 5.1% value for anaemia stated in the MS was an incorrect reference based on a previously-examined scenario. The manufacturer explained regarding the use of the TARGET study "*this is a conservative assumption as the inclusion of AEs from the RECORD-1 study for BSC would have resulted in higher costs for these patients and as a result more favourable ICERs for axitinib*".

Table 5.15 outlines the assumptions made and costs calculated for each of the AEs included in the model.

Table 5.15: List of adverse events and summary of costs included in the economic model (corrected version of MS Table 45)

Adverse event	Study arm and frequency	Cost per episodes	Assumptions
Hypertension	Axitinib arm, 15.3% BSC arm, 2%	£424.00	2 GP visits per year (cost per 11.7 minute visit = £36.00,) 2 district nurse visits per year (cost per visit = £38) Medication for hypertension (cost per year = £276 (inflated to 2011)
Source: [NICE clinical guideline 34]Hypertension medicine "Management of hypertension in adults in primary care: partial update: Costing Report" NICE (2006)[44] http://www.nice.org.uk/nicemedia/pdf/CG034costingreport.pdf Table 2: Future drug costs			
Diarrhoea	Axitinib arm, 10.0%	£544.00	2 days hospitalization
Source: Code VC42Z Rehabilitation for other disorders			
Anaemia NB: This was part of the original MS but corrected to 0% after the clarification letter	BSC arm, 5.1%	£2,068.47	Reported in Mickisch et al 2010, inflated to 2011 costs (PSSRU tariff) [45]

Abbreviations: BSC, best supportive care; GP, general practitioner; PSSRU, Personal Sociak Services Research Unit.

In response to the clarification letter (Section B – Question 22), the manufacturer states that cost estimates for hypertension and diarrhoea were based on the resource use assumed in the PenTAG economic model.[3] These assumptions were further validated with expert clinical opinion to ensure consistency with current clinical practice. They argued that the use of Mickisch et al. 2010[46] to estimate the costs of hypertension and diarrhoea would have a negligible impact on the manufacturers' model results due to relatively small mean AE costs per patient.

Comment

Regarding the lack of systematic search of relevant source data for resource identification in the MS, we refer to section 4.1.1. The ERG considers the cost estimates used in this study valid, though one minor issue was identified.

For the incidence of adverse events in the BSC arm, the manufacturer stated to have used the data from the TARGET study[23]. However, in the TARGET study, the percentage of patients with hypertension was less than 1, and not 2% as applied in the model (see Table 4.5, Section 4.2.1). Given the small costs associated with treating this adverse event, the impact on the ICER is negligible.

Relating to the specification of uncertainty around the cost estimates, two issues were identified.

First, the ERG found an error in the relative dosing intensity (RDI) for the base case scenario. As explained for the utilities in Section 5.2.7, the mean RDI observed in the AXIS trial (102%) is a sample mean and therefore it should be accompanied with a standard error (SE) instead of a SD. Thus, the SD reported in Table 5.13 of 35.2% should be replaced by an SE of 1.86% in the PSA. This is

expected to have a large impact on the uncertainty associated to the costs which is reflected in the new PSA that the ERG performed in Section 5.3.

Second, the cost of death is considered as a fixed (deterministic) amount in the model developed by the manufacturer. These costs were derived from Coyle et al (1999),[41] as a weighted average of the cost of palliative care in hospitals £2,285.88 (N=95, SD=£2,096.80) and in hospices £3,049.91 (N=231, SD=£1,791.70) and inflated from 1999 to 2011 using the PSS Pay & prices inflator. However, the ERG considers that treating these costs as fixed underestimates the uncertainty associated to the costs in the model. Moreover, the paper by Coyle et al (1999) provides all the elements needed to compute the SE. Therefore, the results of the PSA performed by the ERG in Section 5.3 included an SE equal to 104.43 for the cost of death.

5.2.9 Cost effectiveness results

Base Case Analysis

Table 5.16 presents the disaggregated incremental life years gained and QALYs by health state. The incremental life years and incremental QALYs are given by [redacted] for the cytokine refractory subgroup. For the sunitinib refractory population these are [redacted] The disaggregated costs by health state (with and without the PAS) can be seen in Table 5.17. The incremental costs with PAS amount to [redacted] without PAS for the cytokine refractory subgroup. For the sunitinib refractory subgroup these are [redacted] with and without PAS, respectively.

Table 5.16: Summary of QALY accumulation by health state

Health state	LYG (axitinib)	LYG (BSC)	Incremental LYG	QALY (axitinib)	QALY (BSC)	Incremental QALYs
Prior cytokine						
Progression free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior sunitinib						
Progression free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; LYG, Life years gained; QALY, quality-adjusted life year.

Table 5.17: Summary of costs by health state (with and without PAS)

Health state	Cost (axitinib)	Cost (BSC)	Increment
Prior cytokine			
Progression free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]

Health state	Cost (axitinib)	Cost (BSC)	Increment
Total	[redacted]	[redacted]	[redacted]
Prior sunitinib			
Progression free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; LYG, Life years gained.

Table 5.18 and Table 5.19 present the base case results of the cost-effectiveness of axitinib compared with BSC with and without the PAS cost for axitinib, respectively. The additional cost per QALY with PAS amount to £65,326 and to [redacted] without PAS for the cytokine refractory subgroup. For the sunitinib refractory subgroup this is £40,933 with PAS and [redacted] without PAS.

Table 5.18: Base case results – with PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) (QALYs)
Cytokine refractory							
BSC	[redacted]	[redacted]	[redacted]				
Axitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£65,326
Sunitinib refractory							
BSC	[redacted]	[redacted]	[redacted]				
Axitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£40,933

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; STC, simulated treatment comparison.

Table 5.19: Base case results – without PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) (QALYs)
Cytokine refractory							
BSC	[redacted]	[redacted]	[redacted]				
Axitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sunitinib refractory							
BSC	[redacted]	[redacted]	[redacted]				
Axitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; STC, simulated treatment comparison.

Comment

The ERG had to conclude that the base case for the cytokine refractory population performed so far was not optimal, as one of the HR applied was incorrect. For this reason the ERG ran the manufacturers cost-effectiveness model using the PFS hazard ratio for the cytokine refractory population equal to 0.203. The results of these ERG analyses are shown in section 5.3.

5.2.10 Sensitivity analyses

The manufacturer assessed the various uncertainties in the economic evaluation through deterministic sensitivity analysis, scenario analysis and probabilistic sensitivity analysis. While the first two show which parameters and assumption have the largest impact on the model outcomes, the latter shows the overall uncertainty around the ICER. All three type of sensitivity analyses are discussed in the next paragraphs.

5.2.10.1Deterministic sensitivity analyses

Univariate sensitivity analyses were conducted in the MS to test the sensitivity of the results (ICER) to plausible variation of input parameters. Parameter values were varied ±20% to the base case value and the results were displayed in tornado diagrams. In the final version of the MS, sensitivity analyses are presented for the base case with the PAS only and the corresponding tornado diagrams can be seen in Figures 5.8 and 5.9.

Figure 5.8: Base case tornado diagram: Cytokine refractory population (with PAS)

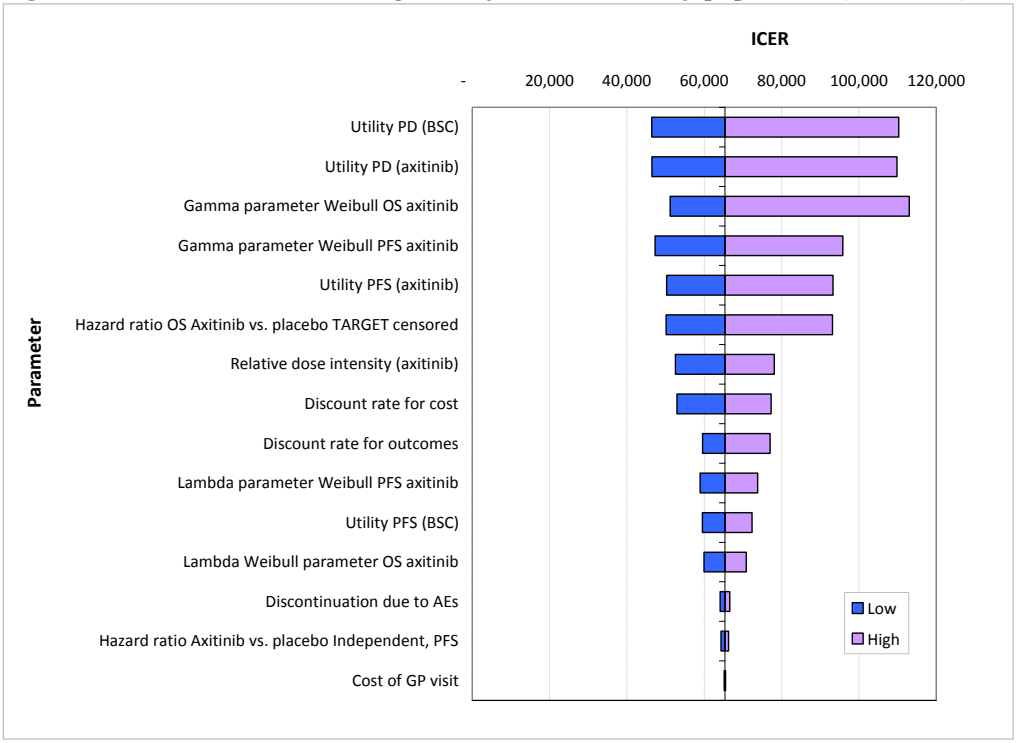
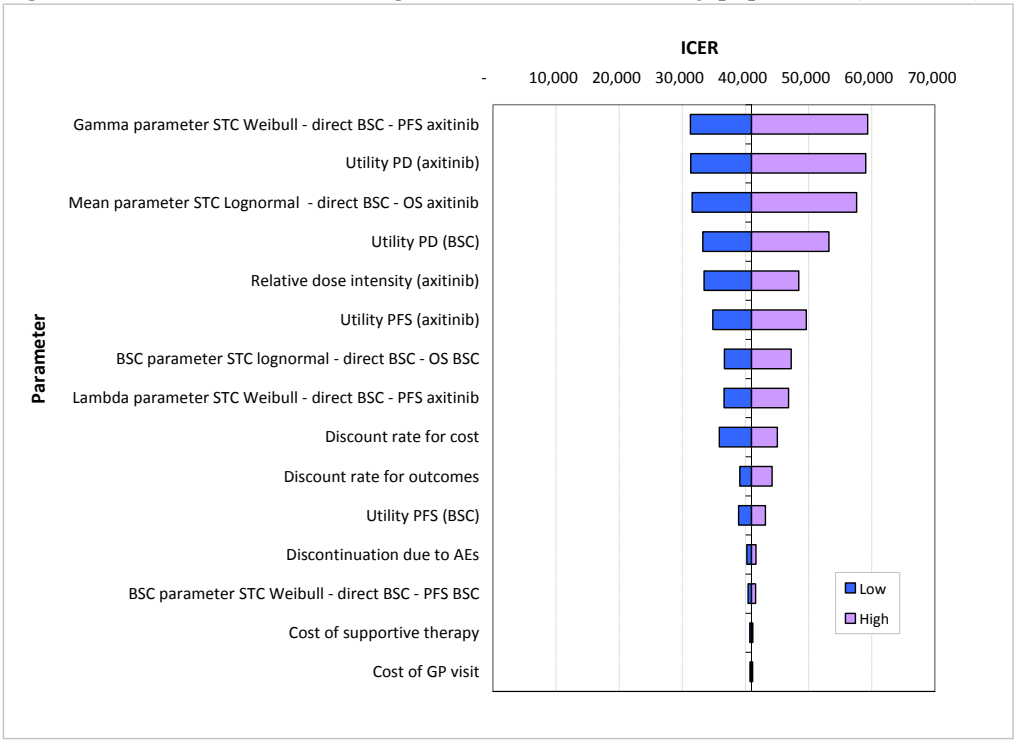


Figure 5.9: Base case tornado diagram – Sunitinib refractory population (with PAS)



From this analysis, the manufacturer concluded that for the cytokine refractory population the cost-effectiveness of axitinib compared to BSC is stable to most changes in the model parameters. The utility parameters (in particular those associated to PD state), survival parameters and the OS hazard ratio of Axitinib versus BSC via the TARGET study indirect comparison had most influence on the model. On the other hand, the cost of a GP visit, the PFS hazard ratio and the discontinuation due to adverse events seem to have little influence on the model results.

For the sunitinib refractory population this analysis suggests that the cost-effectiveness of axitinib compared with BSC is stable to most changes in the model parameters with all variables resulting in upper bound ICERs below £60,000/QALY. The key sources of uncertainty in the model include the survival parameters for PFS and OS, progressed disease utilities, and relative dose intensity of axitinib. Costs (GP visit and supportive therapy) had least influence on the model.

Tornado diagrams for the base case without PAS were shown in the final version of the MS without PAS and can be seen in Figures 5.10 and 5.11.

Figure 5.10: Base case tornado diagram – cytokine refractory population (without PAS)

Figure 5.11: Base case tornado diagram – Sunitinib refractory population (without PAS)

For the cytokine refractory population the ranking of parameters provided by the univariate sensitivity analysis without PAS is very similar to the ranking with PAS, i.e. some of the top parameters have interchanged their positions (since they have similar range of ICER variation) and the bottom nine parameters are exactly the same. For the sunitinib refractory population the two rankings (with and without PAS) are equal. Thus, the same conclusions mentioned above for the situation with PAS also hold here.

In response to the clarification letter (Section B – Question 25), the univariate sensitivity analysis was performed with parameter variation based on 95% confidence intervals (as used in the PSA) for both with and without PAS scenarios. In this case, tornado diagrams 5.12 to 5.15 showed the variation with respect to the incremental net benefit (assuming willingness to pay equal to £50,000). This is because, according to the manufacturer, some scenarios produced negative ICERs, and therefore the incremental net benefit can be better interpreted. From this second analysis, the manufacturer concluded that the utility parameters had most influence on the model, in particular those associated to the progressed disease state.

Figure 5.12: Tornado diagram (INB): sunitinib refractory population (with PAS)

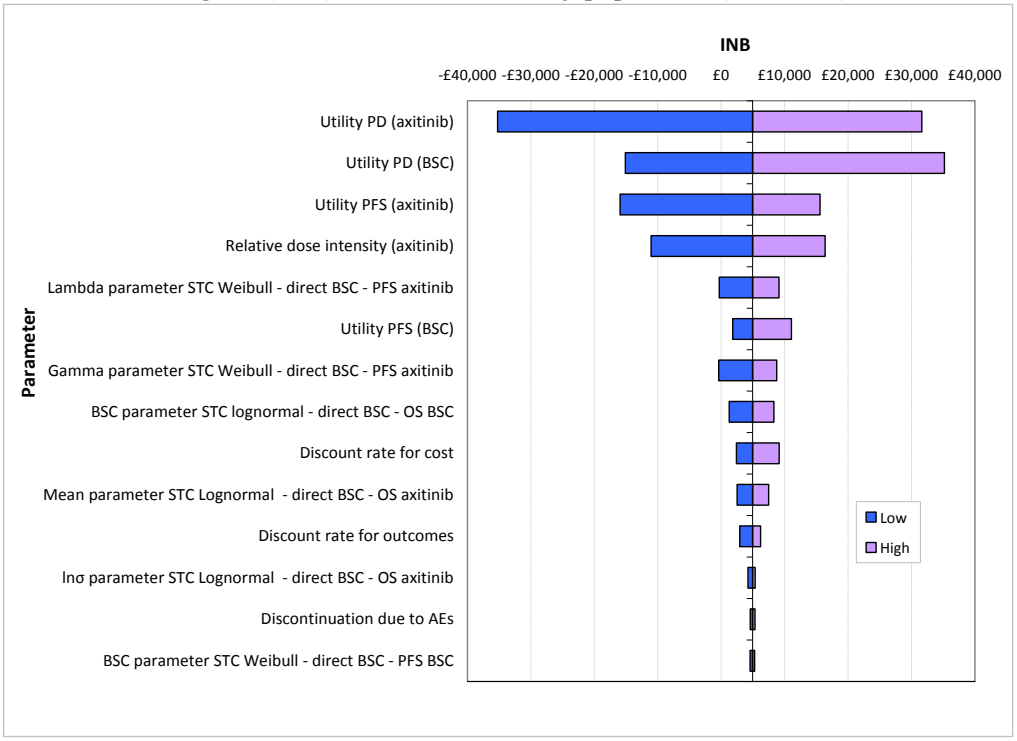


Figure 5.13 (INB): Tornado diagram: cytokine refractory population (with PAS)

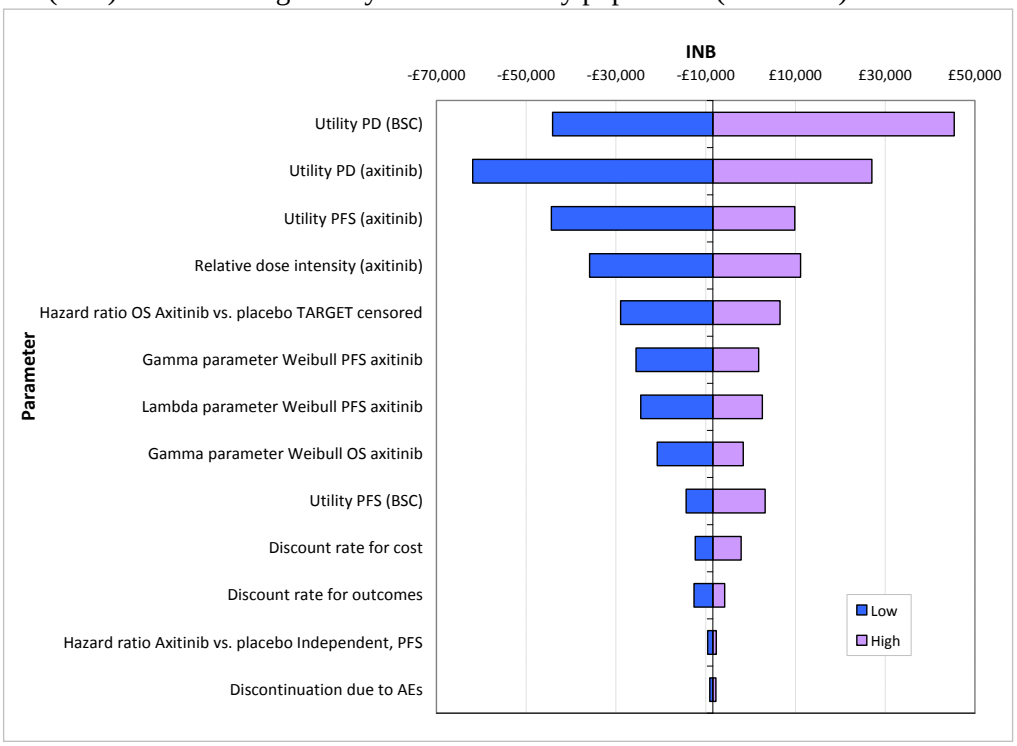


Figure 5.14 (INB): Tornado diagram: sunitinib refractory population (without PAS)

Figure 5.15 (INB): Tornado diagram: cytokine refractory population (without PAS)

5.2.10.2 Scenario analyses

A number of structural assumptions were examined in the manufacturer submission to explore the impact on model outcomes. Specifically, assumptions were tested around the survival distribution chosen to extrapolate axitinib OS and PFS, the method of comparison to best supportive care, utility measurement, dosing intensity, and medical management. The results (ICERs) of the different scenarios (both with and without PAS) are displayed in Table 5.20 for the cytokine refractory population and in Table 5.21 for the sunitinib refractory population.

Table 5.20: Scenario analysis results (with and without PAS) – Cytokine refractory population

Parameter	**Base case**	**Scenario analysis**	**ICER (with and without PAS)**
Base Case			£65,326 [redacted]
Method of PFS extrapolation	Weibull	Lognormal Gompertz	£71,535 [redacted] £63,702 [redacted]
Method of OS extrapolation	Weibull	Loglogistic Gompertz	£52,260 [redacted] 84,255 [redacted]
Axitinib and BSC utility estimates	AXIS study	2nd-line utilities (advanced/mRCC MTA and everolimus appraisal)	£59,654 [redacted]

Axitinib relative dosing intensity	AXIS study	Estimated real-world dosing intensity (Everolimus appraisal)	£51,546 [redacted]
Ongoing medical management in pre-progression state	GP Management	Oncologist Management	£66,410 [redacted]
Time horizon	10 years	5 years 15 years	£83,752 [redacted] £64,359 [redacted]
Discount Rate	3.5% costs and QALYs	0% 6%	£60,015 [redacted] £69,164 [redacted]

Abbreviations: BSC, best supportive care; GP, general practitioner; ICER, incremental cost-effectiveness ratio; mRCC, metastatic renal cell carcinoma; MTA multiple technology appraisal; OS, overall survival; PFS, progression-free survival.

Table 5.21: Scenario analysis results (with and without PAS) – Sunitinib refractory population

Parameter	**Base case**	**Scenario analysis**		**ICER (with and without PAS)**
Base Case				£40,933 [redacted]
Method of PFS comparison	STC Weibull via ITT RECORD-1 BSC population	STC Lognormal via ITT RECORD-1 BSC		£42,428 [redacted]
		STC Weibull via everolimus sunitinib refractory – BSC PFS		£40,509 [redacted]
Method of OS comparison	STC Lognormal via RECORD-1 ITT BSC population	STC Weibull via RECORD-1 ITT BSC		£39,906 [redacted]
		STC Weibull via everolimus sunitinib refractory – BSC RPSFT		£33,268 [redacted]
		RENCOMP	Weibull	£56,113 [redacted]
			Lognormal	£43,384 [redacted]
			Gompertz	£54,851 [redacted]
Axitinib and BSC utility estimates	AXIS study	2nd-line utilities (advanced/mRCC MTA and everolimus appraisal)		£37,059 [redacted]

Parameter	Base case	Scenario analysis	ICER (with and without PAS)
Axitinib relative dosing intensity	AXIS study	Estimated real-world dosing intensity (Everolimus appraisal)	£32,846 [redacted]
Medical management pre-progression	GP Management	Oncologist Management	£42,074 [redacted]
Time horizon	10 years	5 years [redacted] 15 years	£48,283 [redacted] £39,207 [redacted]
Discount Rate	3.5% costs and QALYs	0% [redacted] 6%	£38,254 [redacted] £42,806 [redacted]

Abbreviations: BSC, best supportive care; GP, general practitioner; ICER, incremental cost-effectiveness ratio; mRCC, metastatic renal cell carcinoma; MTA, multiple technology appraisal; OS, overall survival; PFS, progression-free survival; RPSFT, rank preserving structural time failure; STC, simulated treatment comparison.

Based on the above results, the manufacturer concluded that for the sunitinib refractory subgroup, the model is robust to the majority of structural assumptions made. Most of the scenarios examined, including extrapolation assumptions, the STC approach, and source of utility estimate, produced ICERs lower than £50,000/QALY (with PAS). The exception was the method of OS comparison used that produced quite different results ranging from values very close to the base case ICER to a maximum difference larger than £20,000. In particular, the use of the RENCOMP model analysis resulted in ICERs higher than the base case (ranging from between £43,384 and £56,113 with PAS). According to the manufacturer, this indicates that the incremental survival benefit assumed over BSC is a key driver of the model result.

For the cytokine refractory population the manufacturer concluded that the key parameters which increased the ICER included use of a lognormal model to extrapolate axitinib PFS, and a Gompertz model to extrapolate OS (which, as discussed in Section 7.3.2.1 of the MS, are viewed as unrealistic estimates) and use of oncologist management instead of GP management. In any case, none of the scenarios explored by the manufacturer provided an ICER below £50,000. Finally, note that, as mentioned in Section 5.2.5, prolonging the time horizon from 10 to 15 years has a minimal impact on the ICER.

5.2.10.3 Probabilistic sensitivity analyses

A probabilistic sensitivity analysis (PSA) was performed in the manufacturer submission to study the uncertainty around all input parameters of the economic model simultaneously. Probability distributions were specified for all input parameters and cost-effectiveness results associated with simultaneously selecting random values from those distributions were generated. We refer to Table 40 in the MS for all details on distributions and their parameters used for the PSA.

In the final version of the manufacturer's submission, PSA results are presented with PAS only. In this case, axitinib was cost-effective in comparison with BSC for 31% of the observations (out of

1,000 generated in the PSA), i.e. 31% of the observations were below a cost-effectiveness threshold of £50,000 per QALY for the cytokine refractory population (see figures 5.16 and 5.17) whereas this was 67% for the sunitinib refractory population (see figures 5.18 and 5.19). The PSA outcomes plotted in the CE-plane showed how the overall uncertainty is distributed. This was especially large for the cytokine refractory group where approximately 30% of the observations were on the NW quadrant (i.e. they were never cost-effective). For the sunitinib refractory population, although all PSA outcomes were in the NE quadrant, the uncertainty was also large in both costs and effects.

Figure 5.16: Base case cost-effectiveness acceptability curve – Cytokine refractory population (with PAS)

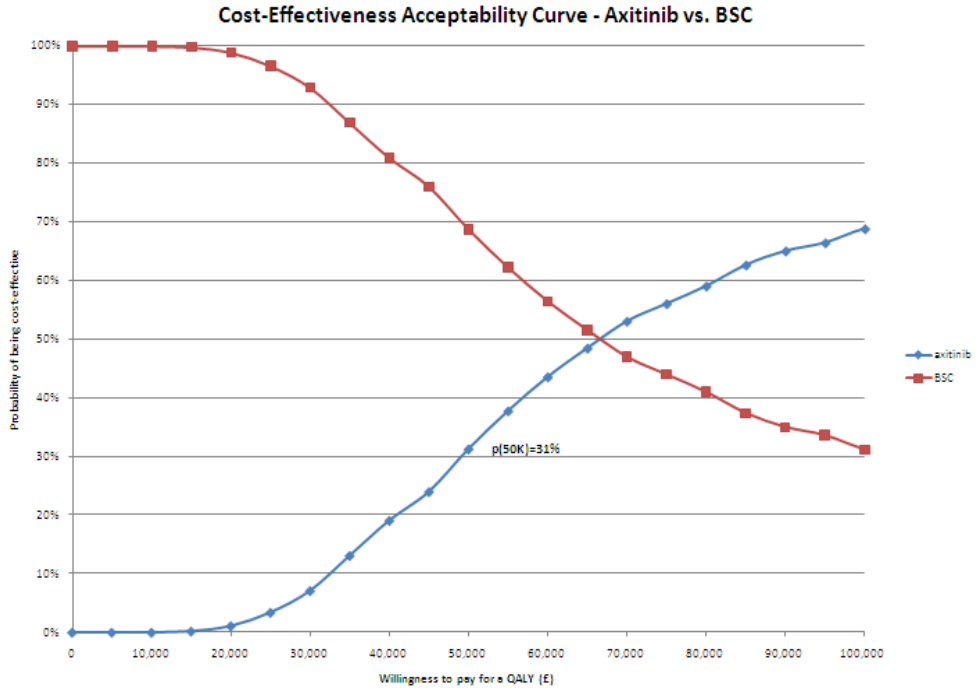


Figure 5.17: Base case PSA scatter plot – Cytokine refractory population (with PAS)

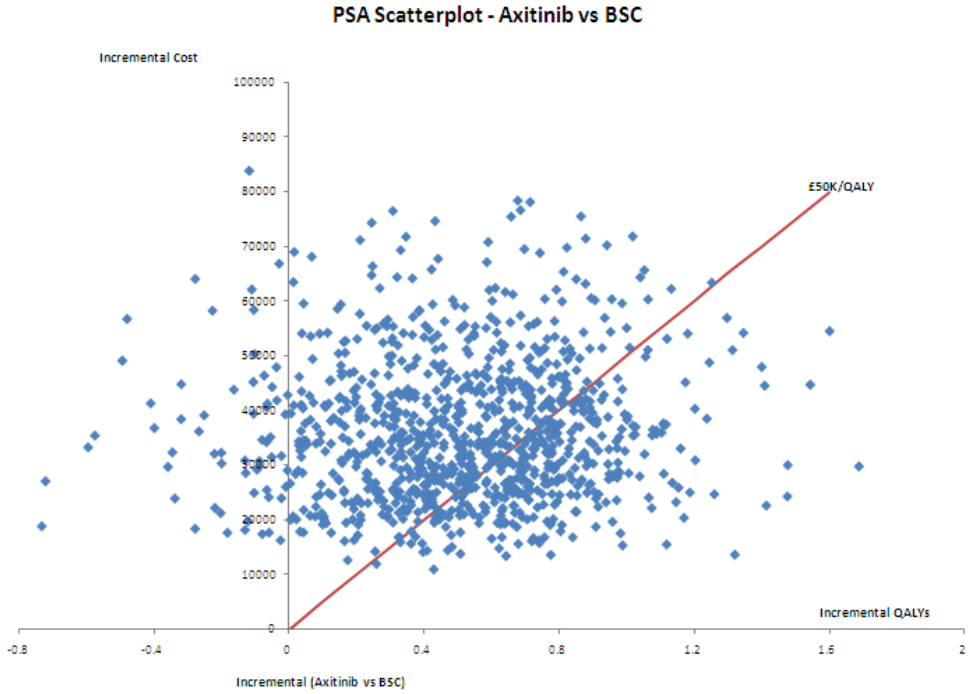


Figure 5.18: Cost-effectiveness acceptability curve – Sunitinib refractory population (with PAS)

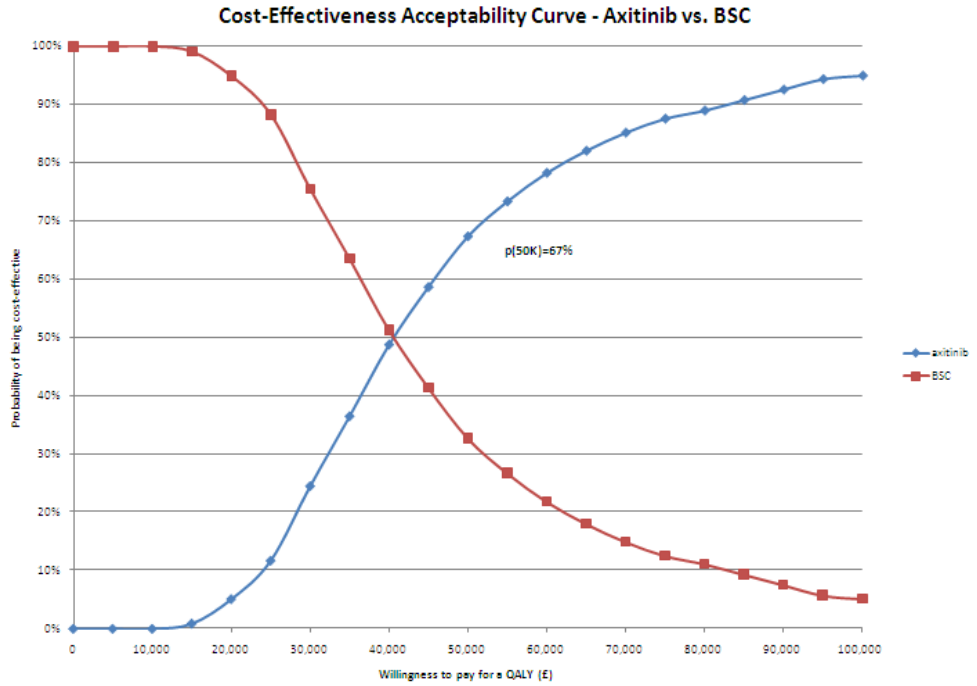
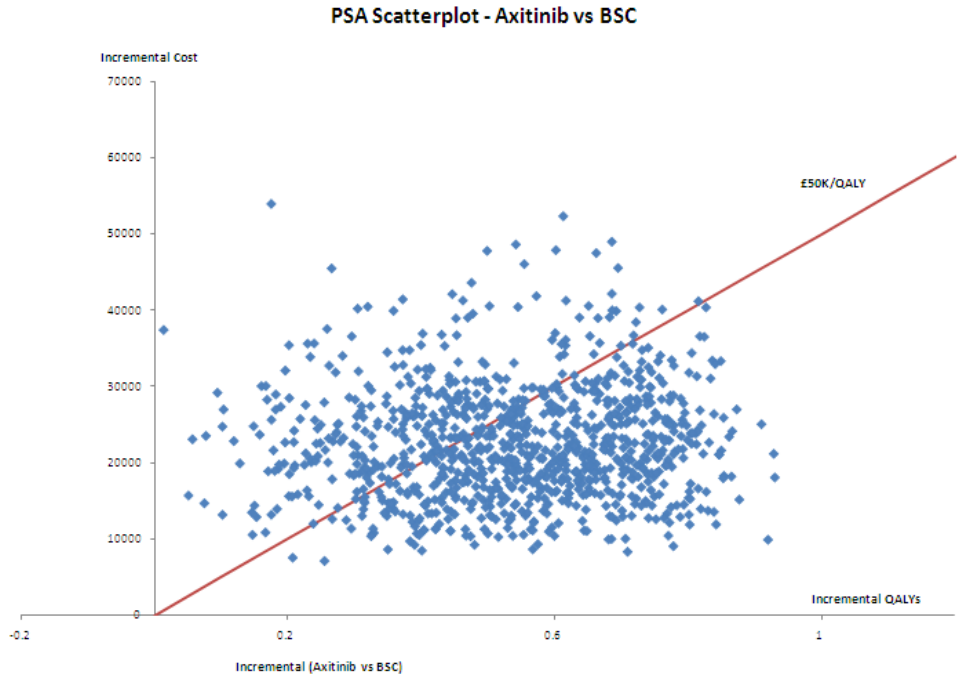


Figure 5.19: PSA scatter plot – Sunitinib refractory population (with PAS)



PSA results without PAS were shown in the final version of the MS without PAS. In this case, the cost-effectiveness probability of axitinib for a threshold of £50,000 per QALY decreased to [redacted] for the cytokine refractory population (see figures 5.20 and 5.21) and to [redacted] for the sunitinib refractory population (sees figures 5.22 and 5.23). The overall uncertainty increased on the costs side and it is also larger for the cytokine refractory group.

Figure 5.20: Cost-effectiveness acceptability curve – Cytokine refractory population (without PAS)

[redacted]

Figure 5.21: PSA scatter plot – Cytokine refractory population (without PAS)

Figure 5.22: Cost-effectiveness acceptability curve – Sunitinib refractory population (without PAS)

Figure 5.23: PSA scatter plot – Sunitinib refractory population (without PAS)

Comment

As mentioned above, the ERG found that the reported PFS hazard ratio for the cytokine refractory population 0.251 with 95% CI (0.165-0.379) should be replaced by 0.203 with 95% CI (0.132, 0.318). This is expected to have a minimal impact on the results of the univariate sensitivity analysis. However, in sections 5.2.7 and 5.2.8.3 it was found that standard deviations were used where standard errors should have been used. Replacing the standard errors for the base case PF and PD utilities, i.e. SD=0.275 for PF and SD=0.316 for PD, by SE=0.0035 and SE=0.0175, and for the relative dosing intensity, i.e. SD=35.2% by SE=1.86%, are expected to have a strong influence on the results of the univariate and probabilistic sensitivity analysis since the uncertainty on both utility and cost sides is substantially decreased.

Based on the scenarios presented in Section 5.2.10.2 the ERG agrees that for the sunitinib refractory subgroup, the model is robust to the majority of structural assumptions made. Moreover, most of the scenarios examined produced ICERs lower than £50,000/QALY (with PAS). However, as mentioned in Section 4.3, the evidence for this subgroup relies on an STC, which is basically a comparison of two individual treatment arms (i.e. this comparison is not based on randomised treatment allocation), for which there is considerable potential for bias in the outcomes of the analysis. Unfortunately, there is no way to assess whether or not the final result will be biased. For the cytokine refractory population the manufacturer concluded that the base case can be viewed as a conservative estimate since most of the ICERs found in the different scenarios are lower than the base case ICER. The ERG does not agree this conclusion since Table 5.20 shows that exactly half of the scenarios are below the base case ICER. Therefore, it remains uncertain whether the base case ICER represents a conservative choice or not. What can be concluded is that for the cytokine refractory population the model does not seem to be as robust to structural assumptions as for the sunitinib refractory subgroup.

Besides the scenarios performed by the manufacturer presented in Section 5.2.10.2, the ERG explored the impact of the subgroup specific utilities in an additional scenario analysis. Moreover, the scenarios for the cytokine refractory population were repeated with the correct PFS hazard ratio.

The ERG want to emphasize that the analyses presented here focus on uncertainty that is quantified, either through a confidence interval around a point estimate or by the definition of scenarios. However, an important source of uncertainty is not part of the sensitivity analyses, and this relates to the STC adjustment factors, which were used to derive PFS and OS curves for the BSC group in the sunitinib refractory subgroup. The manufacture explains in section 7.6 "*The STC adjustment factors are not included in the PSA. Theoretically when the underlying survival curves (i.e. axitinib survival) change, the whole calibration procedure would need to be redone. So the assumption in the model is that while the survival curve parameters for axitinib change as well as the hazard ratios (if applicable), the relationship between the survival curve parameters of axitinib and everolimus or BSC remains constant.*"

This of course leads to a clear, but unquantifiable, under estimation of uncertainty in the model.

Finally, based on the findings mentioned above, the ERG performed probabilistic sensitivity analyses with corrected inputs. The results of all the analyses performed by the ERG can be found in Section 5.3.

5.2.11 Model validation

The manufacturer states in section 7.8 of the MS that a comprehensive and rigorous quality check of the model was performed by a peer-reviewer not involved in the model development. This process involved checking intermediate calculations, implementation, expected function of parameters and the functionality of built-in Macro programs. This repeatable process produced a checklist spreadsheet indicating the specific tasks performed, and their results returned. In the clarification letter (Section B – Question 26) the manufacturer was asked to submit such spreadsheet to the ERG.

The manufacturer also provided a table comparing the model results with the clinical data. This table (Table 5.22) can be used as an indicator of the internal validity of the model since it shows the median PFS and OS values for axitinib in the AXIS study by subgroup (with a 95% confidence interval) and the model base case estimates. It can be observed that all median estimates are within the 95% confidence intervals of the AXIS trial estimates. According to the manufacturer, "*these results demonstrate that the modelled figures are comparable to the clinical trial results observed*".

Table 5.22: Summary of model results for axitinib compared with clinical data

Outcome	**Clinical trial result (months, median)**	**Model result (months, median)**
Prior cytokine		
PFS	12.1 (10.1-13.9)	11.6
OS	29.4 (24.5-NE)	33.3
Prior sunitinib		
PFS	4.8 (4.5-6.4)	6.32
OS	15.2 (12.8-18.3)	16.6

Abbreviations: NE, not estimable; OS, overall survival; PFS, progression-free survival.

The MS points out the absolute survival estimate produced by the model for treatment with BSC as the key source of uncertainty in the model (see MS Section 7.10.1). The base case analysis performed by the manufacturer produced a BSC median survival estimate of 8.3 months, using the STC via the ITT BSC population and a lognormal extrapolation. Moreover, the everolimus appraisal resulted in median BSC overall survival estimates of between 8.9 and 10.8 months which was judged by clinical

and patient opinion as an overestimate of the true BSC overall survival. This was further supported by the systematic review carried out to examine BSC survival post-sunitinib failure (see MS Section 6.7.10) where the majority of the estimates found were in the four to six month range, substantially lower than the RECORD-1 estimate (10 months). Furthermore, this result was consistent with the 5.8 months median OS observed in the BSC arm of the RENCOMP study. These results can be seen as a gauge of the face validity of the model.

Comment

The spreadsheet attached to the clarification letter is not of the form of a checklist as the ERG had expected, i.e. a predefined list of tests to be performed with expected outcomes, to thoroughly test the technical integrity of the model. Instead of that, it shows a post-testing list of bugs found in the model and the corresponding action to fix them. Thus, the ERG cannot assess the quality of the validation process from the model validation spread sheet, i.e. it is not clear how extensively the model was reviewed.

The ERG is of the opinion that the fact that the median estimates predicted by the model are within the 95% confidence intervals of the AXIS trial estimates does not ensure that the model results are comparable to those observed in the clinical trial results. Results could have been more comparable if model results were also accompanied by their corresponding 95% confidence interval. Medians could be judged to be similar in magnitude but this would be a subjective assessment since no formal test for the difference in median was provided. However, using this subjective assessment, the model seems to reasonably reproduce the trial results.

Regarding the face validity of the BSC results, the comparisons made by the manufacturer seem to indicate that for most studies, BSC OS was smaller than that found in this study. This would indeed imply that the current estimates for BSC are conservative. But given the higher median OS found in the everolimus appraisal,[47] it cannot be stated with certainty that the current estimate is conservative.

5.3 Additional work undertaken by the ERG

5.3.1 ERG base case

As mentioned in Section 4.2.6, an error was found in the calculation of the PFS hazard ratio for the cytokine refractory population. Thus, the values shown in Table 5.8 should be replaced by 0.203 with 95% CI equal to (0.132, 0.318). Moreover, as mentioned in Section 5.2.7 and in Section 5.2.8.1, the SD for the base case utilities used in Table 5.11 (0.275 and 0.316) and the SD for the base case relative dosing intensity used in Table 5.13 (35.2%) do not reflect the true parameter uncertainty since the standard errors (0.0035 and 0.0175 for the utilities and 1.86% for the relative dosing intensity) have to be considered. Moreover, as explained in Section 5.2.8, the cost of death should be considered uncertain with a mean value equal to £3,923 and SE equal to 104.43. With these new values the ERG defined a new ERG base case and performed the subsequent analyses. For the deterministic analyses (base case and scenarios defined by the manufacturer), only the cytokine refractory group was affected by the change made, thus the results of the sunitinib refractory group remains unchanged and can be found in Table 5.18 (base case with PAS), Table 5.19 (base case without PAS) and Table 5.19 (scenarios with and without PAS).

Base case cytokine refractory group

As can be seen in the table below, the effect of replacing the PFS hazard ratio is minimal since the ICERs obtained are approximately [redacted] lower than those reported by the manufacturer and shown in Table 5.18 and Table 5.19.

Table 5.23: ERG base case results – cytokine refractory group (with and without PAS)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) (QALYs)
Cytokine refractory group							
BSC	[redacted]	[redacted]	[redacted]				
Axitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£64,388 [redacted]

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; STC, simulated treatment comparison.

ERG univariate sensitivity analysis (cytokine refractory and sunitinib refractory subgroups)

The tornado plots presented in figure 5.24 to 5.27 are markedly different from those presented by the manufacturer (see figure 5.8 to 5.15). This is due to the fact that here all parameters have been varied between the lower and upper limit of their 95% confidence interval, rather than between plus and minus 20%. This is especially clear in the cytokine refractory subgroup, where varying the HR for the OS of axitinib versus BSC leads to an ICER up to £400,000 when the upper limit is 0.99.

Figure 5.24: ERG univariate sensitivity analysis – Sunitinib refractory population (with PAS)

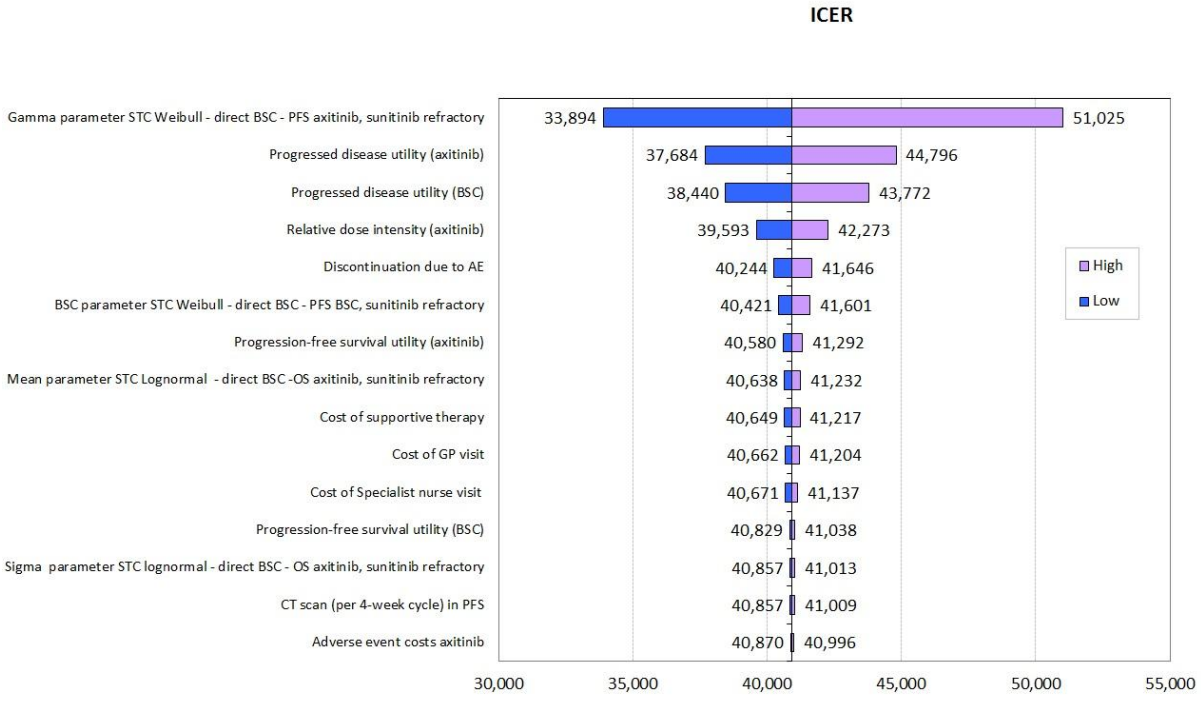


Figure 5.25: ERG univariate sensitivity analysis – Sunitinib refractory population (without PAS)

Figure 5.26: ERG univariate sensitivity analysis – Cytokine refractory population (with PAS)

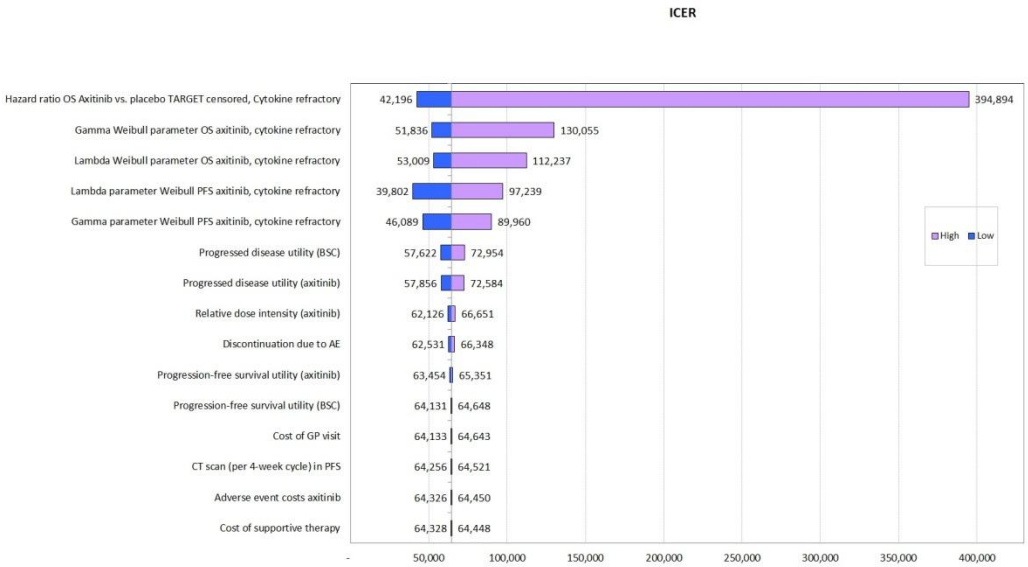


Figure 5.27: ERG univariate sensitivity analysis – Cytokine refractory population (without PAS)

Scenario analysis cytokine refractory group

As for the base case, the ICERs for the various manufacturer defined scenarios obtained by replacing the PFS hazard ratio (see Table 5.21) are approximately [redacted] lower than those reported by the manufacturer and shown in Table 5.20.

Table 5.24: ERG scenario analysis results (with and without PAS) – Cytokine refractory population

Parameter	Base case	Scenario analysis	ICER (with and without PAS)
Base Case			£64,388 [redacted]
Method of PFS extrapolation	Weibull	Lognormal Gompertz	£70,637 [redacted] £62,736 [redacted]
Method of OS extrapolation	Weibull	Loglogistic Gompertz	£51,618 [redacted] £82,793 [redacted]
Axitinib and BSC utility estimates	AXIS study	2nd-line utilities (advanced/mRCC MTA and everolimus appraisal)	£58,886 [redacted]
Axitinib relative dosing intensity	AXIS study	Estimated real-world dosing intensity (Everolimus appraisal)	£50,733 [redacted]

Ongoing medical management in pre-progression state	GP Management	Oncologist Management	£65,462 [redacted]
Time horizon	10 years	5 years 15 years	£82,284 [redacted] £63,446 [redacted]
Discount Rate	3.5% costs and QALYs	0% 6%	£59,216 [redacted] £68,119 [redacted]

Abbreviations: BSC, best supportive care; GP, general practitioner; ICER, incremental cost-effectiveness ratio; mRCC, metastatic renal cell carcinoma; MTA multiple technology appraisal; OS, overall survival; PFS, progression-free survival.

Probabilistic Sensitivity Analysis (cytokine refractory and sunitinib refractory subgroups)

PSA results are first presented with PAS. In this case, for a cost-effectiveness threshold of £50,000 per QALY, axitinib was cost-effective for 19% of the observations (out of 1,000 generated in the PSA) for the cytokine refractory population and 83% for the sunitinib refractory population, respectively (see figures 5.28 and 5.30 below). Note that the manufacturer reported that axitinib was cost-effective for 31% of the observations for the cytokine refractory subgroup and for 67% of the observations for the sunitinib refractory population.

As mentioned in the beginning of Section 5.3, the base case utilities will experience a reduction in their uncertainty since their SD (0.275 and 0.316) has been replaced by the SE (0.0035 and 0.0175). This is expected to reduce the PSA uncertainty along the x-axis. On the costs side, the uncertainty in the relative dosing intensity will be also reduced since the SD (35.2%) has also been replaced by the SE (1.86%). However, the inclusion of the cost of death as an uncertain parameter will increase the uncertainty associated to the costs. The overall effect (increase or decrease) on the uncertainty along the y-axis will depend on the magnitude of the uncertainty associated to the RDI and the cost of death.

The PSA outcomes plotted in the CE-plane (see figures 5.29 and 5.31) illustrate this effect for the cytokine and sunitinib refractory subgroups, respectively.

Compared to the manufacturer's base case with PAS in Figure 5.17, it is clear that now the uncertainty on the costs side has been reduced since most of the PSA outcomes are comprised between £30,000 and £40,000 on the y-axis, whereas in the manufacturer's base case with PAS, the PSA outcomes ranged from £10,000 to £80,000 on the y-axis. Therefore, it seems clear that the uncertainty reduction associated to the RDI outweighs the increase in uncertainty associated to the cost of death since the overall uncertainty on the costs side is reduced with respect to the manufacturer's base case. Regarding the effects (x-axis), the PSA outcomes are still spread over the NE quadrant along the x-axis direction. However, only less than 1% of the observations (in contrast to approximately 30% reported by the manufacturer) are now in the NW quadrant which clearly indicates a reduction of the uncertainty on the effects side as well.

For the sunitinib refractory population the uncertainty has been dramatically reduced compared to the manufacturer's base case shown in Figure 5.19 and all the PSA outcomes seem to converge towards

the ICER. However, as mentioned in Section 4.6.2 and at the end of Section 5.2.10, the assumptions made by the manufacturer for the STC underestimate in a way that cannot be quantified the overall uncertainty of the model. Nevertheless, given the result of the ERG PSA, we may think that the underestimation is large since the uncertainty estimated for the sunitinib refractory subgroup is much smaller than for the cytokine refractory population. This seems to be counterintuitive because for the cytokine refractory population there is at least a (limited) source of evidence which does not exist for the sunitinib refractory group. For that reason, the ERG base case scenario for the sunitinib refractory population must be interpreted with caution; in fact, it can be seen as a best case scenario or lower limit for the overall uncertainty associated to the model for the sunitinib refractory subgroup.

Figure 5.28: ERG cost-effectiveness acceptability curve – Cytokine refractory population (with PAS)

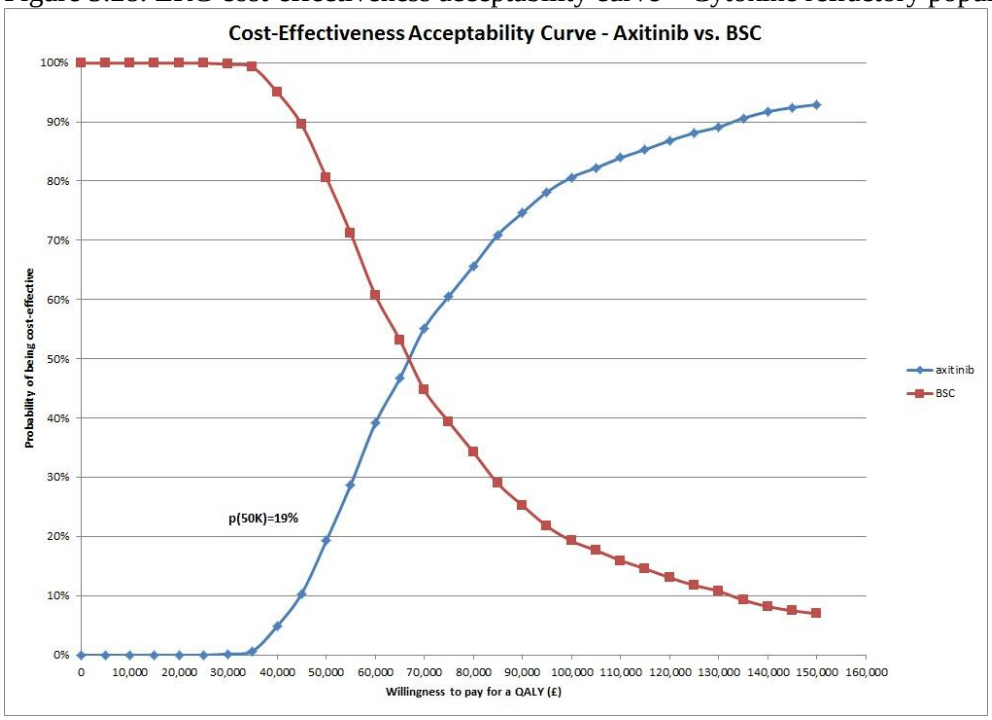


Figure 5.29: ERG PSA scatter plot – Cytokine refractory population (with PAS)

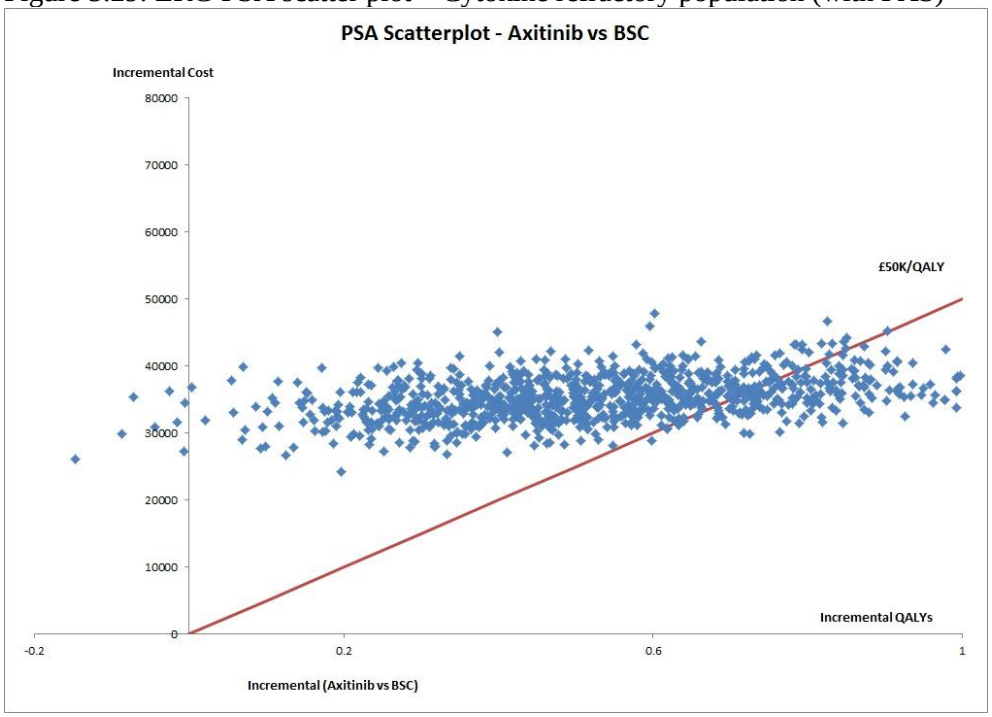


Figure 5.30: ERG Cost-effectiveness acceptability curve – Sunitinib refractory population (with PAS)

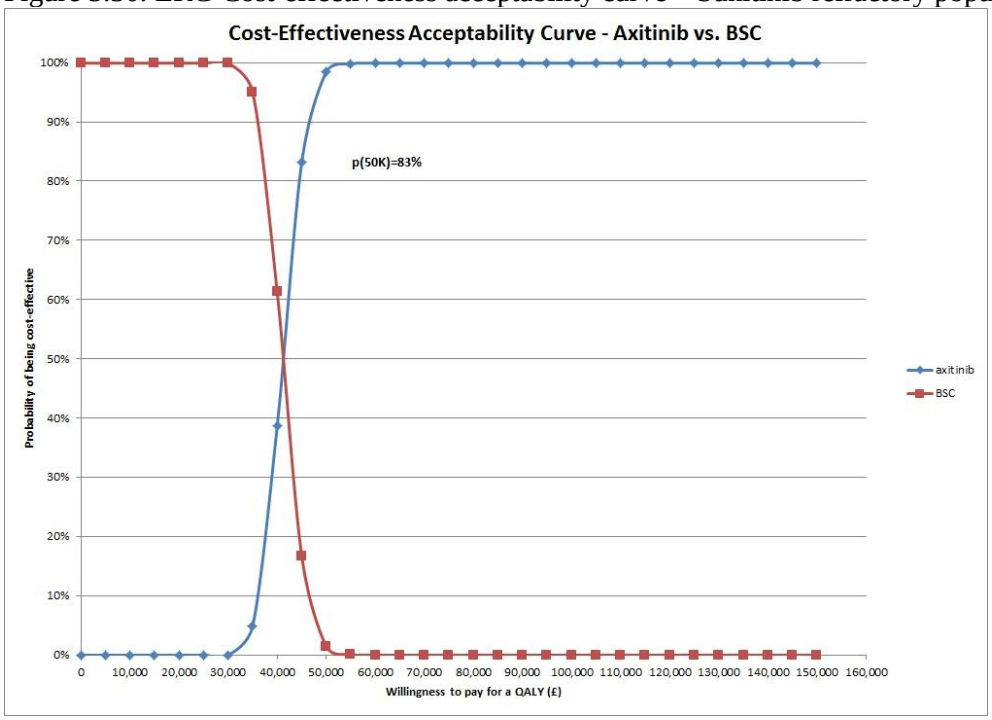
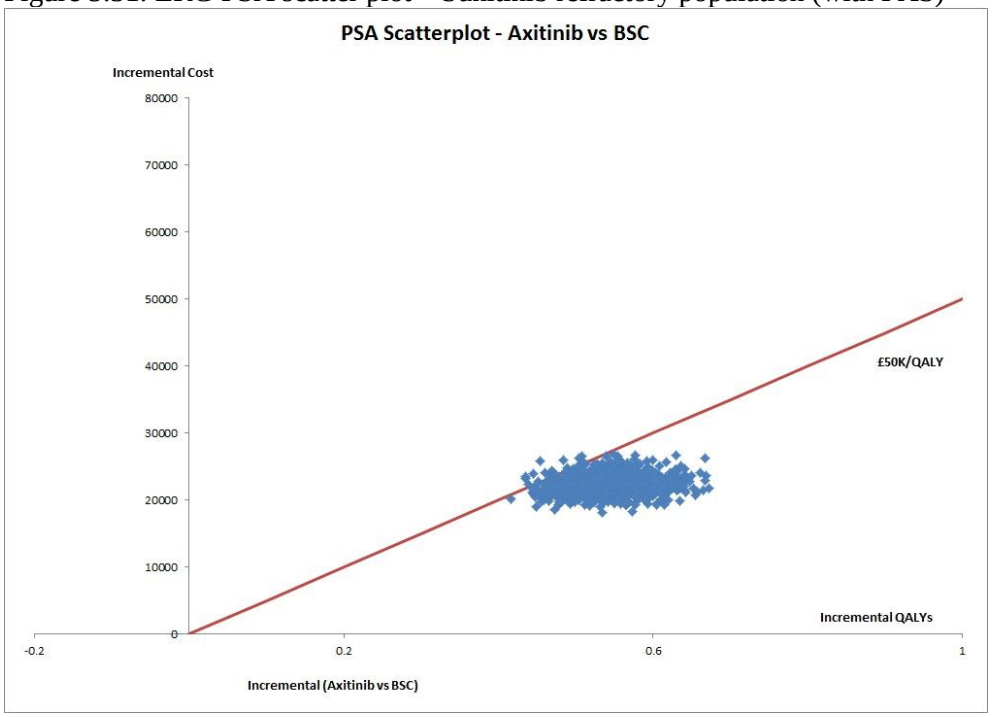


Figure 5.31: ERG PSA scatter plot – Sunitinib refractory population (with PAS)



PSA results without PAS can be seen below. The cost-effectiveness probability of axitinib for a threshold of £50,000 per QALY was [redacted] for the cytokine refractory population (see figures 5.32 and 5.33) and [redacted] for the sunitinib refractory population, respectively (sees figures 5.34 and 5.35). Note that the manufacturer reported that axitinib was cost-effective for [redacted] of the observations for the cytokine refractory subgroup and for [redacted] of the observations for the sunitinib refractory population. Thus, in both cases axitinib is significantly less cost-effective which clearly illustrates the effect of increasing the price of axitinib, i.e. when axitinib costs are higher the PSA outcomes tend move up into the CE-plane. Therefore, more PSA outcomes are found above the threshold of £50,000 per QALY. Similar remarks to those made for the situation with PAS regarding the overall distribution of the uncertainty also apply in this case.

Figure 5.32: ERG Cost-effectiveness acceptability curve – Cytokine refractory population (without PAS)

Figure 5.33: ERG PSA scatter plot – Cytokine refractory population (without PAS)

Figure 5.34: ERG Cost-effectiveness acceptability curve – Sunitinib refractory population (without PAS)

Figure 5.35: ERG PSA scatter plot – Sunitinib refractory population (without PAS)

5.3.2 Additional utility scenarios

Subgroup specific utilities scenario

In response to the clarification letter (Section B – Question 15), the manufacturer provided the utility estimates for the progression free and progressive disease health state for the prior sunitinib and the prior cytokine subgroups separately. These are reported in Table 5.12 in Section 5.2.7.

The ICERs obtained using subgroup specific utilities can be seen in Table 5.25. Note that they are similar to those obtained in the base case (using the overall utility value) but lower ICERs are obtained for the cytokine refractory population whereas the opposite is observed for the sunitinib refractory group. Both scenarios further assumed, as in the base case, that there is no difference in utilities between the axitinib and the BSC arms.

Table 5. 25: Incremental results with and without PAS using subgroup specific utilities.

ICER (£) incremental (QALYs)	Base-case (no PAS)	Subgroup specific utilities (no PAS)	Base case (PAS)	Subgroup specific utilities (PAS)
Prior Cytokine	[redacted]	[redacted]	£64,388	£62,885
Prior Sunitinib	[redacted]	[redacted]	£40,933	£42,095

Lower utilities scenario

In section 5.2.7, it was discussed that the EQ-5D health states as measured in patients during the AXIS trial were valued using a US tariff developed by Shaw et al.[38] Since studies have shown that the US valuation is consistently higher than the UK valuation[39,40] the ERG has assessed the impact of lowering the current utilities (PF 0.692, PD 0.61). Based on table 3 from the paper by Johnson et al. [40] we lower the 0.69 utility value to 0.66, and the 0.61 utility value to 0.54.

The ICERs obtained in this case can be seen in Table 5.26. Note that these are slightly higher than those obtained in the base case (around [redacted] higher without PAS and £3,000 with PAS) for both populations. Both scenarios further assumed, as in the base case, that there is no difference in utilities between the axitinib and the BSC arms.

Table 5.26: Incremental results with and without PAS using lower utilities.

ICER (£) incremental (QALYs)	Base-case (no PAS)	Lower utilities (no PAS)	Base case (PAS)	Lower utilities (PAS)
Prior Cytokine	[redacted]	[redacted]	£64,388	£67,123
Prior Sunitinib	[redacted]	[redacted]	£40,933	£44,125

Higher BSC progression-free utility scenario

As the utility data was collected in the AXIS trial, no data was available in the BSC group. In section 5.2.7 it was discussed that it would be assumed that the utility estimate for axitinib and BSC is equal. While valid arguments were given, it is of course possible that due to for example side effects, the quality of life in the intervention group is a bit higher than in the BSC group. To test to impact of

treatment specific utility estimates before progression, we assumed that the utility in BSC would be a bit higher, i.e. 0.72 instead of the 0.69 used for axitinib.

Using these values we find that this higher utility for BSC has a very minimal impact on the ICERs (see Table 5.27).

Table 5.27: Incremental results with and without PAS using higher progression-free utility for BSC

ICER (£) incremental (QALYs)	Base-case (no PAS)	Lower axitinib PF utility (no PAS)	Base case (PAS)	Lower axitinib PF utility (PAS)
Prior Cytokine	[redacted]	[redacted]	£64,388	£65,385
Prior Sunitinib	[redacted]	[redacted]	£40,933	£41,363

5.4 Conclusions

The economic model described in the MS is considered by the ERG to meet the NICE reference case[5] and is in-line with the decision problem specified in the scope.

The ERG assessment indicated that the model was generally well presented and reported. Some minor issues were identified regarding the use of standard deviations in the PSA instead of standard errors; whilst these errors have a clear impact on the outcomes of the univariate and probabilistic sensitivity analyses, they do not impact the estimate of the central ICER. In addition, the clinical assessment revealed an error in the hazard ratio for the BSC arm in the cytokine refractory subgroup. This error was corrected but this only had a minimal impact on the ICER.

The ERG univariate sensitivity analysis revealed that for the cytokine refractory subgroup, the ICER is extremely sensitive to changes in the HR for the overall survival. At the upper limit of the 95% confidence interval, the ICER would amount to almost £400,000 (with PAS). This is in sharp contrast with the results for the sunitinib refractory subgroup, where changes to input parameters lead to modest changes in the ICER. This is related to the fact that no measures of uncertainty were provided for the adjustment factor for the BSC arm, which means that this uncertainty was not considered in the univariate (or probabilistic) sensitivity analysis.

The cost-effectiveness results were generally robust under the scenario analyses conducted, though a few scenarios impacted the ICER considerably. For the cytokine refractory subgroup, the most important assumption relates to the extrapolation of the OS in the axitinib arm, when a Gompertz distribution is used rather than a Weibull distribution, the ICER increases sharply. For the sunitinib refractory subgroup, the factor having a significant impact was the approach used to model OS in the BSC arm; using RENCOMP observational data with an indirect comparison led to a substantially higher ICER.

It is important to realize is that many uncertainties related to the health economic evaluation in the sunitinib refractory subgroup have not been quantified, and thus are not represented in the central estimates of the ICER or in the CEACs. As discussed in chapter 4, both the STC approach and the RENCOMP approach to estimating the PFS and OS of the BSC group have the potential for considerable bias, either upwards or downwards. This means that the same is true for the ICERs reported in chapter 5. In addition, the STC approach lacks an estimation of the uncertainty

surrounding the point estimates it provides. Again, this also means that the uncertainty around the sunitinib-refractory ICER is most likely severely underestimated.

For the cytokine refractory subgroup most uncertainties have been taken into account, revealing a large uncertainty in the number of QALYs gained, and thus around the ICER.

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

None of the additional clinical and economic analysis undertaken by the ERG resulted in central ICERs that varied from the manufacturers results in any meaningful way.

7 END OF LIFE

Where appropriate, this section should summarise the manufacturer's case for using the NICE end of life treatment criteria and discuss to what extent the manufacturer's argument is valid.

NICE has issued supplementary advice to the Appraisal Committees for appraising life-extending, end of life treatments. These are treatments which may be life-extending for patients with short life expectancy, and which are licensed for indications affecting small numbers of patients with incurable illnesses.[48]

The criteria for appraisal of end of life treatments are described by NICE as follows:

2.1 This supplementary advice should be applied in the following circumstances and when all the criteria referred to below are satisfied:

2.1.1 The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;

2.1.2 There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment, and;

2.1.3 The treatment is licensed or otherwise indicated, for small patient populations. (Source: NICE Supplementary advice to the Appraisal Committees, page 2)

Table 58 of the manufacturer's submission (MS, page 180, see Table 7.1 below) outlines the end of life criteria used for axitinib in the manufacturer's submission and Pfizer's justification for applying the end of life criteria to axitinib (Source: Pfizer Submission, Section 7.10.1, P180).

Table 7.1: End of life criteria for axitinib

Criteria	**Justification**
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	All model cases examined for sunitinib refractory patient population result in mean BSC survival estimates of less than 24 months. In addition, the systematic review of survival after sunitinib failure carried out to support this submission indicates that real-world survival times in absence of second-line treatment are expected to be less than a year.
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment	Axitinib results in expected survival gains of greater than 3 months over BSC in all model cases evaluated.
The treatment is licensed or otherwise indicated, for small patient populations	The annual number of patients eligible to receive axitinib in the sunitinib or cytokine refractory patient population is 1580 in year 1, rising to 1743 in year 5.

According to Pfizer, *"there are approximately 8,163 incident kidney cancers in England and Wales every year and RCC accounts for 90% of all kidney cancers. Of these patients, 27% and 14% are expected to have stage III and IV (advanced/metastatic (m)RCC) disease, respectively, and 33% of former stage I-II are expected to recur to stage III-IV, resulting in approximately 4456 patients*

diagnosed with advanced/mRCC per year (NICE TA169)[8]." (Source: Pfizer submission, Executive Summary, p11)

In their appraisal of sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma the NICE appraisal committee stated that it "was aware that the total number of people with advanced and/or metastatic RCC in England and Wales was approximately 4000." (Source: NICE, TA169[8], p.22) Furthermore, the committee considered that for appraisal TA169[8], sunitinib should be regarded as meeting the population criterion for an end-of-life treatment.

In addition, in their appraisal of everolimus for the treatment of second line advanced RCC the NICE appraisal committee stated that it "was aware that in England and Wales the total number of people who would be eligible for treatment with everolimus was less than 4000. The Committee heard from the clinical specialist that the life expectancy for people with advanced RCC receiving best supportive care alone was unlikely to be greater than 24 months and was potentially as low as five months." (Source: NICE, TA219[35], p.30) . In their conclusion, the Committee was satisfied that everolimus met the criteria for being a life-extending, end-of-life treatment, and that the evidence presented for this consideration was sufficiently robust (Source: NICE, TA219[35], p.30).

There is no direct evidence to suggest that axitinib increases survival by more than three months compared with best supportive care. However, in the AXIS trial "median overall survival in the axitinib arm was 15.2 months (95% CI: 12.8-18.3) in the sunitinib refractory subgroup and 29.4 months (95% CI: 24.5-NR) in the cytokine refractory group" (*Source: Pfizer submission, Section 6, p33*). Using data from the TARGET trial (cytokine refractory population), the median actuarial overall survival among patients in the placebo group was 14.7 months.[23]

As mentioned before, an indirect comparison in the sunitinib refractory population is not possible. However, for the cytokine-refractory population an indirect comparison can be performed using data from the AXIS and TARGET trials. For the cytokine-refractory population in the AXIS trial, axitinib has a median overall survival of 29.4 months versus 27.8 months for sorafenib, an advantage of 1.6 months in favour of axitinib. The TARGET trial, when using data uncensored for cross-over, showed a median overall survival of 17.8 months for sorafenib versus 15.2 months for placebo, an advantage of 2.6 months in favour of axitinib. When using data censored for cross-over the TARGET trial shows a median overall survival of 17.8 months for sorafenib versus 14.3 months for placebo, an advantage of 3.5 months in favour of axitinib. Overall, for the cytokine-refractory population, the advantage for axitinib over placebo is 4.2 months when using data uncensored for cross-over from the TARGET trial and 5.1 months when using data censored for cross-over from the TARGET trial. (Source: Pfizer submission, Table 16, page 93)

Therefore, the ERG agrees that axitinib meets the end of life criteria as specified by NICE.

8 CONCLUSIONS

Overall, the manufacturer's submission is clearly presented and the analyses and underlying assumptions are sound and reasonable. The main issue with this submission is whether a simulated treatment comparison (STC) presents a valid and reliable estimate of the clinical effectiveness of axitinib versus BSC in a sunitinib refractory population. As there is no direct trial evidence it is not possible to compare the results of the STC to any existing evidence so the accuracy and reliability of the results cannot be ascertained. In addition, the uncertainty around the STC results is missing.

The ERG found one error in the indirect comparison (cytokine refractory population) but this did not affect the conclusions. The methods for the STC (sunitinib refractory population) follow the recommended method and it appears to be well-conducted and clearly explained, including considerations of the limitations of the method and available trial evidence. However, the STC is basically a comparison of individual treatment arms and is open to considerable bias. There is no way to assess whether or not the final result is biased.

In addition, the reliability of the results of the included studies and the comparability of the trials included are potential issues. These have been adequately reported in the manufacturer's submissions, and have been summarised in this ERG report.

The economic model described in the MS is considered by the ERG to meet the NICE reference case. The model structure was considered to be appropriate and the ERG has no major concerns regarding the selection of data used within the model, beyond the issues relating to the estimation of the treatment effectiveness of axitinib versus BSC.

In conclusion, the ICER for axitinib in a cytokine refractory population is based on a well performed indirect comparison. The results seem reliable, and present an ICER with large confidence intervals due to considerable uncertainty regarding the treatment effect of axitinib on overall survival. The ICER for axitinib in a sunitinib refractory population is based on an STC which did not included the uncertainty surrounding the estimates used in the analyses (indicating that the uncertainty around the ICER is only a lower limit of the true uncertainty), and is basically a comparison of individual treatment arms and is therefore open to considerable bias.

8.1 Implications for research

The NICE scope specified BSC as the only comparator for axitinib in adult patients with advanced renal cell carcinoma (aRCC) after failure of prior treatment with sunitinib or a cytokine. For the cytokine refractory population there is sufficient evidence to allow an indirect comparison between axitinib and BSC. For the sunitinib refractory population, there is not a network to link axitinib with BSC. Therefore a randomised trial comparing axitinib with BSC would be the first research priority. Alternatively, treatments that have been compared with axitinib in a sunitinib refractory population, such as sorafenib and temsirolimus, could also be compared with BSC, to allow an indirect comparison between axitinib and BSC.

9 REFERENCES

[1] National Institute for Health and Clinical Excellence. Single Technology Appraisal: axitinib for the treatment of advanced renal cell carcinoma after failure of prior systemic treatment: final scope [Internet]. London: NICE, 2012 [cited 19.9.12] Available from: **http://www.nice.org.uk/nicemedia/live/13688/59699/59699.pdf**

[2] Pfizer Ltd. Axitinib for the treatment of advanced renal cell carcinoma after failure of prior systemic treatment: Submission to National Institute of Health and Clinical Excellence. Single Technology Appraisal (STA). London: Pfizer Ltd., 2012: 400.

[3] Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, et al. Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation [Internet]. Exeter: Peninsula Technology Assessment Group (PenTAG), 2008 [cited 27.9.12] Available from: **http://www.nice.org.uk/nicemedia/live/11817/41488/41488.pdf**

[4] Pitt M, Crathorne L, Moxham T, Bond M, Hyde C. *Everolimus for the second-line treatment of advanced and/or metastatic renal cell carcinoma [Internet]*. Exeter: Peninsula Technology Assessment Group (PenTAG), 2009 [cited 24.5.12] Available from: **http://www.nice.org.uk/nicemedia/live/12044/47366/47366.pdf**

[5] National Institute for Health and Clinical Excellence. *Guide to the methods of technology appraisal [Internet]*. London: NICE, 2008 [cited 11.10.12] Available from: **http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf**

[6] Cancer Research UK. Kidney cancer incidence statistics [Internet]. London: Cancer Research UK, 2012 [cited 18.9.12]. Available from: **http://www.cancerresearchuk.org/cancer-info/cancerstats/types/kidney/incidence/#source1**

[7] National Institute for Health and Clinical Excellence. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma. Technology Appraisal 169: costing template [Internet]. London: NICE, 2009 [cited 19.9.12] Available from: **http://guidance.nice.org.uk/TA169/CostingTemplate/xls/English**

[8] National Institute for Health and Clinical Excellence. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma. NICE technology appraisal guidance 169 [Internet]. London: NICE, 2009 [cited 24.5.12] Available from: **http://www.nice.org.uk/nicemedia/live/12143/43556/43556.pdf**

[9] The British Association of Urological Surgeons (BAUS). BAUS cancer registry analyses of minimum data set for urological cancers: January 1st – 31st December 2008 [Internet]. London: Section of Oncology, BAUS, 2009 [cited 18.9.12]. Available from: **http://www.baus.org.uk/Resources/BAUS/Documents/PDF%20Documents/Data%20and%20Audit/2008finalanalyses.pdf**

[10] National Institute for Health and Clinical Excellence. Bevacizumab (first-line), sorafenib (first- and second-line), sunitinib (second-line) and temsirolimus (first-line) for the treatment of advanced and/or metastatic renal cell carcinoma. NICE technology appraisal guidance 178 [Internet]. London: NICE, 2009 [cited 24.5.12] Available from: **http://www.nice.org.uk/nicemedia/live/12220/45232/45232.pdf**

[11] Poffiri E, Miscoria M, Lim L, Shamash J, Chowdhury S, Powles T. The outcome of patients who fail sunitinib and do not have access to sunitinib therapy [Abstract 931P]. Paper presented at 35th ESMO Congress; 8–12 October 2010; Milan, Italy *Ann Oncol* 2010;21(Suppl. 8):viii292.

[12] Miscoria M, Pirrie S, Baijal S, Tew A, James ND, Porfiri E. Analysis of survival after disease progression in patients with renal cell carcinoma (RCC) who failed treatment with sunitinib [Abstract e15154]. Paper presented at 2011 ASCO Annual Meeting; 3-7 June 2011; Chicago, Illinois *J Clin Oncol* 2011;29(Suppl.15).

[13] Motzer RJ, Bacik J, Schwartz LH, Reuter V, Russo P, Marion S, et al. Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma. *J Clin Oncol* 2004;22(3):454-63.

[14] Motzer RJ, Bander NH, Nanus DM. Renal-cell carcinoma. *N Engl J Med* 1996;335(12):865-75.

[15] National Institute for Health and Clinical Excellence. *Pazopanib for the first-line treatment of advanced renal cell carcinoma. NICE technology appraisal guidance 215 [Internet]*. London: NICE, 2011 [cited 24.5.12] Available from: **http://www.nice.org.uk/nicemedia/live/13346/53185/53185.pdf**

[16] National Institute for Health and Clinical Excellence. Single Technology Appraisal (STA) for everolimus (afinitor®) in advanced renal cell carcinoma: specification for manufacturer/sponsor submission of evidence [Internet]. London: NICE, n.d. [cited 19.9.12] Available from: **http://www.nice.org.uk/nicemedia/live/12044/47345/47345.pdf**

[17] McGowan J, Sampson M, Lefebvre C. An evidence based checklist for the peer review of electronic search strategies (PRESS EBC). *Evidence Based Library and Information Practice* 2010;5(1):1-6.

[18] National Institute for Health and Clinical Excellence. *Single Technology Appraisal: specification for manufacturer/sponsor submission of evidence*. London: NICE, October 2009, 2009. 76p.

[19] Pfizer Ltd. Axitinib for the treatment of advanced renal cell carcinoma after failure of prior systemic treatment – Response to request for clarification from the ERG. London: Pfizer Ltd., 2012: 16.

[20] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [cited 23.03.11] Available from: **http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm**

[21] Pfizer Ltd. Temsirolimus versus sorafenib as second-line therapy in patients with advanced RCC who have failed first-line sunitinib (INTORSECT). NCT00474786. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2012 [cited 26.9.12]. Available from: **http://clinicaltrials.gov/show/NCT00474786**

[22] Rini BI, Escudier B, Tomczak P, Kaprin A, Szczylik C, Hutson TE, et al. Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. *Lancet* 2011;378(9807):1931-9.

[23] Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Siebels M, et al. Sorafenib in advanced clear-cell renal-cell carcinoma. *N Engl J Med* 2007;356(2):125-34.

[24] Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Staehler M, et al. Sorafenib for treatment of renal cell carcinoma: final efficacy and safety results of the phase III treatment approaches in renal cancer global evaluation trial. *J Clin Oncol* 2009;27(20):3312-8.

[25] Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, et al. Phase 3 trial of everolimus for metastatic renal cell carcinoma: final results and analysis of prognostic factors. *Cancer* 2010;116(18):4256-65.

[26] Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, et al. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *Lancet* 2008;372(9637):449-56.

[27] Caro JJ, Ishak KJ. No head-to-head trial? simulate the missing arms. *Pharmacoeconomics* 2010;28(10):957-67.

[28] Di Lorenzo G, Casciano R, Malangone E, Buonerba C, Sherman S, Willet J, et al. An adjusted indirect comparison of everolimus and sorafenib therapy in sunitinib-refractory metastatic renal cell carcinoma patients using repeated matched samples. *Expert Opin Pharmacother* 2011;12(10):1491-7.

[29] Harmenberg U, Lundstam S, Wahlgren T, Kowalski J, Jakobsson M, Sandin R, et al. Treatment and overall survival in metastatic renal cell carcinoma: a swedish population-based study (2000–2008). Paper presented at the 2012 American Society of Clinical Oncology Genitourinary Cancers Symposium; 2-4 Feb 2012; San Francisco, CA. 2012.

[30] Wahlgren T, Kowalski J, Lundstam S, Harmenberg U, Sandström P, Jakobsson M, et al. Evolution of overall survival in renal cell carcinoma (2000–2008): results from a Swedish population-based study. Paper presented at the 2011 European Multidisciplinary Cancer Congress; 23-27 Sept 2011; Stockholm, Sweden. 2011.

[31] Bracarda S, Hutson TE, Porta C, Figlin RA, Calvo E, Grunwald V, et al. Everolimus in metastatic renal cell carcinoma patients intolerant to previous VEGFr-TKI therapy: a RECORD-1 subgroup analysis. *Br J Cancer* 2012;106(9):1475-80.

[32] National Institute for Health and Clinical Excellence. Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy. NICE technology appraisal guidance 171 [Internet]. London: NICE, 2009 [cited 26.9.12] Available from: **http://www.nice.org.uk/nicemedia/live/11898/44812/44812.pdf**

[33] Hoyle M, Green C, Thompson-Coon J, Liu Z, Welch K, Moxham T, et al. Cost-effectiveness of sorafenib for second-line treatment of advanced renal cell carcinoma. *Value Health* 2010;13(1):55-60.

[34] Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8(36):iii-iv, ix-xi, 1-158.

[35] National Institute for Health and Clinical Excellence. *Everolimus for the second-line treatment of advanced renal cell carcinoma. NICE technology appraisal guidance 219 [Internet]*. London: NICE, 2011 [cited 24.5.12] Available from: **http://www.nice.org.uk/nicemedia/live/13437/54111/54111.pdf**

[36] Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin* 2010;26(5):1091-6.

[37] Zbrozek AS, Hudes G, Levy D, Strahs A, Berkenblit A, DeMarinis R, et al. Q-TWiST analysis of patients receiving temsirolimus or interferon alpha for treatment of advanced renal cell carcinoma. *Pharmacoeconomics* 2010;28(7):577-84.

[38] Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005;43(3):203-20.

[39] Kharroubi SA, O'Hagan A, Brazier JE. A comparison of United States and United Kingdom EQ-5D health states valuations using a nonparametric bayesian method. *Stat Med* 2010;29(15):1622-34.

[40] Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? *Med Care* 2005;43(3):221-8.

[41] Coyle D, Small N, Ashworth A, Hennessy S, Jenkins-Clarke S, Mannion R, et al. Costs of palliative care in the community, in hospitals and in hospices in the UK. *Crit Rev Oncol Hematol* 1999;32(2):71-85.

[42] Curtis L. *Unit costs of health and social care 2011 [Internet]*. Canterbury: Personal Social Services Research Unit, 2011 [cited 27.9.12] Available from: **http://www.pssru.ac.uk/uc/uc2011contents.htm**

[43] Department of Health. *NHS reference costs 2010-2011 [Internet]*. London: Department of Health, 2011 [cited 27.9.12] Available from: **http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140**

[44] National Institute for Health and Clinical Excellence. *Management of hypertension in adults in primary care: partial update. NICE clinical guideline 34: costing report [Internet]*. London: NICE, 2006 [cited 27.9.12] Available from: **http://www.nice.org.uk/nicemedia/pdf/CG034costingreport.pdf**

[45] Mickisch G, Gore M, Escudier B, Procopio G, Walzer S, Nuijten M. Costs of managing adverse events in the treatment of first-line metastatic renal cell carcinoma: bevacizumab in combination with interferon-alpha2a compared with sunitinib. *Br J Cancer* 2010;102(1):80-6.

[46] Mickisch G, Gore M, Escudier B, Procopio G, Walzer S, Nuijten M. Costs of managing adverse events in the treatment of first-line metastatic renal cell carcinoma: bevacizumab in combination with interferon-alpha2a compared with sunitinib. *Br J Cancer* 2010;102(1):80-6.

[47] National Institute for Health and Clinical Excellence. Advice on everolimus for the second-line treatment of advanced renal cell carcinoma: decision of the panel. London: NICE, 2011

[48] National Institute for Health and Clinical Excellence. *Appraising treatments which may extend life, at the end of life: supplementary advice to the appraisal committees [Internet]*. London: NICE, 2009 [cited 19.9.12] Available from: **http://www.nice.org.uk/media/E4A/79/SupplementaryAdviceTACEoL.pdf**

[49] InterTASC Information Specialists' Sub-Group (ISSG). *ISSG Search Filter Resource [Internet]*. York: Centre for Reviews and Dissemination, University of York, 2011 [cited 26.9.12] Available from: **http://www.york.ac.uk/inst/crd/intertasc/**

[50] Aiello E, Muszbek N, Richardet E, Lingua A, Charbonneau C, Remak E. Cost-effectiveness of new targeted therapy sunitinib malate as second line treatment in metastatic renal cell carcinoma in Argentina (Abstract PCN15). Paper presented at ISPOR 2007.

[51] Casciano R, Chulikavit M, Di Lorenzo G, Liu Z, Baladi J-F, Wang X, et al. Economic evaluation of everolimus versus sorafenib for the treatment of metastatic renal cell carcinoma after failure of first-line sunitinib. *Value Health* 2011;14(6):846-51.

[52] Contreras-Hernandez I, Mould-Quevedo J, Salinas-Escudero G, Tapia-Valencia J, Davila-Loaiza G, Garduño-Espinosa J. A cost-utility analysis model for the second line treatment of metastatic renal cell carcinoma in Mexico (Abstract PCN30). Paper presented at ISPOR 2007.

[53] El Ouagari K CM. Cost-effectiveness of treating metastatic renal cell carcinoma (mRCC) patients whose disease failed on VEGF-TKI therapies with everolimus compared to treating with best supportive care (BSC) alone: a canadian societal perspective [Abstract 927P]. Paper presented at 35th ESMO Congress; 8-12 Oct 2010; Milan, Italy *Ann Oncol* 2010;21(Suppl 8):viii291.

[54] Gao X, Reddy P, Dhanda R, Gondek K, Yeh Y, Stadler W, et al. Cost-effectiveness of sorafenib versus best supportive care in advanced renal cell carcinoma. Paper presented at 2006 ASCO Annual Meeting; 2-6 June 2006; Atlanta, Georgia. *J Clin Oncol* 2006;24(Suppl 18):Abstract 4604.

[55] Gao X, Shah S, Reddy P, Gondek K, Pashos C. Economic evaluation of sorafenib versus best supportive care in advanced renal cell carcinoma: an updated cost-effectiveness analysis [Abstract PCN24]. Paper presented at ISPOR 2008.

[56] Gao X, Reddy P, Dhanda R, Gondek K, Yeh YC, Stadler WM, et al. Cost-effectiveness of sorafenib versus best supportive care in advanced renal cell carcinoma. *Journal of Clinical Oncology. Conference: ASCO Annual Meeting* 2006;24(18S):4604.

[57] Jaszewski B, Gao X, Reddy P, Bhardwaj T, Bjarnason G, Finelli A, et al. Cost effectiveness of sorafenib versus best supportive care in advanced renal cell carcinoma in Canada. Paper presented at 2007 ASCO Annual Meeting; 1-5 June 2007; Chicago, Illinoi. *J Clin Oncol* 2007;25(Suppl 18):Abstract 5111.

[58] Ondrackova B, Demlova R. Sorafenib and sunitinib in metastatic renal cell carcinoma: cost-effectiveness analysis in reimbursement proceedings vs. data from clinical practice. Paper presented at ISPOR 13th Annual European Congress; 6-9 Nov 2010; Prague, Czech Republic *Value Health* 2010;13(7):A267.

[59] Paz-Ares L, del Muro JG, Grande E, Diaz S. A cost-effectiveness analysis of sunitinib in patients with metastatic renal cell carcinoma intolerant to or experiencing disease progression on immunotherapy: perspective of the Spanish National Health System. *J Clin Pharm Ther* 2010;35(4):429-38.

[60] Pitt M, Crathorne L, Moxham T, Bond M, Hyde C. Everolimus for the second-line treatment of advanced and/or metastatic renal cell cancer: a critique of the submission from Novartis. *Health Technol Assess* 2010;14(Suppl. 2):41-6.

[61] Purmonen T, Vuorinen R, Lapela M, Pyrhonen S, Kellokumpu-Lehtinen P. Cost and survival analysis of interferon treatment in metastatic renal cell carcinoma [Abstract 649]. Paper presented at 34th Congress of the European Society for Medical Oncology (ESMO); 12-16 Sept 2008; Stockholm, Sweden *Ann Oncol* 2008;19(Suppl 8):viii206-viii7.

[62] Tatar M, Akbulut H. Cost-effectiveness of sorafenib in unresectable and/or metastatic renal cell carcinoma in Turkey. Paper presented at ISPOR 12th Annual European Congress; 24-27 Oct 2009; Paris, France. *Value Health* 2009;12(7):A222.

[63] Teich V, Fernandes R, Schiola A. Cost-effectiveness analysis of sorafenib associated to best supportive care (BSC) versus best supportive care alone in the second line treatment of advanced renal cell carcinoma under the Brazilian public health care system perspective. Paper presented at ISPOR 14th Annual International Meeting; 16-20 May 2009; Orlando, FL United States. *Value Health* 2009;12(3):A42.

[64] Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, et al. Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation. *Health Technol Assess* 2010;14(2):1-184.

[65] Van Nooten F, Dewilde S, Van Belle S, Marbaix S. Cost-effectiveness of sunitinib as second line treatment in patients with metastatic renal cancer in Belgium [Abstract PCN22.] Paper presented at ISPOR 2007.

Appendix 1A: ERG Search Strategies

The ERG undertook the following search as a focused update to supplement the MS 10.15. The ERG searches were designed to identify RCTs looking at the use of sorafenib, everolimus and temsirolimus in a sunitinib refractory population. For completeness the ERG also added additional synonyms such as alternative drug names, as well as CAS registry numbers in order to retrieve any papers that may have been missed by the original MS searches

Search Strategies

Search A: (Sunitinib AND Sorafenib AND mRCC AND RCTs) NOT (MS search 10.15)
Medline search: Lines #1-52 replicate the MS search. The MS search was run in March 2012, so a date facet was inserted lines #53-55 in order to ensure any new papers retrieved by the original MS strategy would not be removed from the new results set. Lines #56-80 contain the new ERG focused strategy intended to identify any RCTs containing Sunitinib AND Sorafenib AND mRCC. The MS search is then "NOT"-ed from the ERG search in lines 81 in order to remove papers already retrieved by the previous MS search, leaving only new or previously missed references.

M**edline (OvidSP): 1946-2012/8/wk4**
Searched 3.9.12

1 exp Carcinoma, Renal Cell/ (19997)
2 exp Kidney Neoplasms/ (54168)
3 ((renal or kidney) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).mp. (61123)
4 (sunitinib or su?10398 or su ?10398 or sutent).mp. (2142)
5 (hypernephroma* or nephroid carcinoma* or hypernephroid carcinoma*).mp. (1269)
6 Randomized controlled trials as Topic/ (82579)
7 Randomized controlled trial/ (334926)
8 Random allocation/ (75561)
9 Double blind method/ (116639)
10 Single blind method/ (16602)
11 Clinical trial/ (472808)
12 exp Clinical Trials as Topic/ (259930)
13 or/6-12 (834727)
14 (clinic$ adj trial$1).tw. (173771)
15 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw. (114261)
16 Placebos/ (31248)
17 Placebo$.tw. (138522)
18 Randomly allocated.tw. (13922)
19 (allocated adj2 random).tw. (674)
20 or/14-19 (354497)
21 13 or 20 (944356)
22 Case report.tw. (168540)
23 Letter/ (757777)
24 Historical article/ (285776)
25 Review of reported cases.pt. (0)
26 Review, multicase.pt. (0)
27 or/22-26 (1201593)
28 21 not 27 (917618)
29 Epidemiologic studies/ (5478)
30 exp case control studies/ (569366)

31 exp cohort studies/ (1203431)
32 Case control.tw. (62146)
33 (cohort adj (study or studies)).tw. (63137)
34 Cohort analy$.tw. (2808)
35 (Follow up adj (study or studies)).tw. (33332)
36 (observational adj (study or studies)).tw. (31796)
37 Longitudinal.tw. (112083)
38 Retrospective.tw. (217383)
39 Cross sectional.tw. (126617)
40 Cross-sectional studies/ (145750)
41 or/29-40 (1611992)
42 28 or 41 (2336377)
43 1 or 2 or 3 or 5 (63528)
44 4 and 42 and 43 (483)
45 (progress* or fail*).mp. (1384950)
46 (interleukin 2 or bioleukin or IL-2).mp. (66906)
47 ((alpha adj2 interferon) or alferon or cilferon or kemron or veldona).mp. (34524)
48 cytokine*.mp. (233731)
49 4 or 46 or 47 or 48 (306361)
50 45 and 49 (45211)
51 4 or 50 (46549)
52 42 and 43 and 51 (955)
53 2012$.ed. (540733)
54 52 and 53 (63)
55 52 not 54 (892)
56 (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4).af. (2173)
57 (sorafenib or bay 43 9006 or bay 43-9006 or bay 439006 or bay43 9006 or bay43-9006 or bay439006 or nexavar or 284461-73-0).af. (2440)
58 56 and 57 (829)
59 carcinoma, renal cell/ or Kidney Neoplasms/ (51850)
60 ((kidney or renal) adj3 (carcinoma$ or cancer$ or adenocarcinoma$ or neoplasm$ or tumo?r$ or malignan$ or metastas$)).ti,ab,ot,hw. (62087)
61 ((nephroid or hypernephroid) adj3 carcinoma$).ti,ab,ot,hw. (112)
62 (collecting duct adj3 carcinoma$).ti,ab,ot,hw. (226)
63 (hypernephroma$ or RCC or mRCC or (tumor adj2 grawitz)).ti,ab,ot,hw. (8089)
64 exp neoplasm metastasis/ or metastas$.ab,ti. (262646)
65 exp kidney/ or (renal or kidney).ab,ti. (588115)
66 64 and 65 (12862)
67 or/59-63,66 (66440)
68 58 and 67 (496)
69 randomized controlled trial.pt. (334926)
70 controlled clinical trial.pt. (84930)
71 randomized.ab. (238154)
72 placebo.ab. (133998)
73 drug therapy.fs. (1562332)
74 randomly.ab. (171301)
75 trial.ab. (246741)
76 groups.ab. (1122504)
77 or/69-76 (2905862)
78 animals/ not (animals/ and humans/) (3680958)
79 77 not 78 (2468358)
80 68 and 79 (441)
81 80 not 55 (200)

Based on Trials filter:
Lefebvre C, Manheimer E, Glanville J. Chapter 6: searching for studies. Box 6.4.c: Cochrane Highly sensitive search strategy for identifying randomized controlled trials in Medline: Sensitivity-maximizing version (2008 version); OVID format. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org

Medline In-Process & Other Non-Indexed Citations (OvidSP): up to 2012/8/31
Medline Daily Update (OvidSP): up to 2012/8/31
Searched 3.9.12

1 (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4).af. (283)
2 (sorafenib or bay 43 9006 or bay 43-9006 or bay 439006 or bay43 9006 or bay43-9006 or bay439006 or nexavar or 284461-73-0).af. (334)
3 1 and 2 (95)
4 carcinoma, renal cell/ or Kidney Neoplasms/ (21)
5 ((kidney or renal) adj3 (carcinoma$ or cancer$ or adenocarcinoma$ or neoplasm$ or tumo?r$ or malignan$ or metastas$)).ti,ab,ot,hw. (2052)
6 ((nephroid or hypernephroid) adj3 carcinoma$).ti,ab,ot,hw. (0)
7 (collecting duct adj3 carcinoma$).ti,ab,ot,hw. (15)
8 (hypernephroma$ or RCC or mRCC or (tumor adj2 grawitz)).ti,ab,ot,hw. (620)
9 exp neoplasm metastasis/ or metastas$.ab,ti. (9042)
10 exp kidney/ or (renal or kidney).ab,ti. (16465)
11 9 and 10 (496)
12 or/4-8,11 (2324)
13 3 and 12 (67)
14 randomized controlled trial.pt. (770)
15 controlled clinical trial.pt. (71)
16 randomized.ab. (12735)
17 placebo.ab. (5010)
18 drug therapy.fs. (1288)
19 randomly.ab. (12415)
20 trial.ab. (13428)
21 groups.ab. (72810)
22 or/14-21 (96725)
23 animals/ not (animals/ and humans/) (2207)
24 22 not 23 (96321)
25 13 and 24 (25)

Based on Trials filter:
Lefebvre C, Manheimer E, Glanville J. Chapter 6: searching for studies. Box 6.4.c: Cochrane Highly sensitive search strategy for identifying randomized controlled trials in Medline: Sensitivity-maximizing version (2008 version); OVID format. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org

Embase (OvidSP): 1974-2012/wk35
Searched 3.9.12

1 exp kidney carcinoma/ (39053)
2 exp kidney tumor/ (84645)

3 ((renal or kidney) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).mp. (90485)
4 (hypernephroma* or nephroid carcinoma* or hypernephroid carcinoma*).mp. (1508)
5 exp sunitinib/ (9234)
6 (sunitinib or su?10398 or su ?10398 or sutent).mp. (9504)
7 Clinical trial/ (875054)
8 Randomized controlled trial/ (330593)
9 Randomization/ (59294)
10 Single blind procedure/ (16335)
11 Double blind procedure/ (113064)
12 Crossover procedure/ (34836)
13 Placebo/ (216576)
14 Randomi?ed controlled trial$.tw. (78390)
15 Rct.tw. (9935)
16 Random allocation.tw. (1215)
17 Randomly allocated.tw. (17843)
18 Allocated randomly.tw. (1848)
19 (allocated adj2 random).tw. (792)
20 Single blind$.tw. (12773)
21 Double blind$.tw. (136713)
22 ((treble or triple) adj blind$).tw. (304)
23 Placebo$.tw. (184088)
24 Prospective study/ (212616)
25 or/7-24 (1290811)
26 Case study/ (16848)
27 Case report.tw. (239874)
28 Abstract report/ or letter/ (861436)
29 or/26-28 (1113428)
30 25 not 29 (1255043)
31 Clinical study/ (86264)
32 Case control study/ (70056)
33 Family study/ (9694)
34 Longitudinal study/ (55051)
35 Retrospective study/ (290790)
36 Prospective study/ (212616)
37 Randomized controlled trials/ (19708)
38 36 not 37 (212104)
39 Cohort analysis/ (129696)
40 (Cohort adj (study or studies)).mp. (87484)
41 (Case control adj (study or studies)).tw. (63467)
42 (follow up adj (study or studies)).tw. (42598)
43 (observational adj (study or studies)).tw. (46939)
44 (epidemiologic$ adj (study or studies)).tw. (68066)
45 (cross sectional adj (study or studies)).tw. (64024)
46 or/31-35,38-45 (1012257)
47 1 or 2 or 3 or 4 (102826)
48 30 or 46 (2001176)
49 5 or 6 (9504)
50 47 and 48 and 49 (2088)
51 (progress* or fail*).mp. (1866022)
52 (interleukin 2 or bioleukin or IL-2).mp. (97658)
53 ((alpha adj2 interferon) or alferon or cilferon or kemron or veldona).mp. (57445)
54 cytokine*.mp. (353364)
55 49 or 52 or 53 or 54 (460334)
56 51 and 55 (73843)

57 49 or 56 (79367)
58 47 and 48 and 57 (2835)
59 2012$.em. (879230)
60 58 and 59 (190)
61 58 not 60 (2645)
62 sunitinib/ or (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4).mp,rn. (9516)
63 Sorafenib/ or (sorafenib or bay 43 9006 or bay 43-9006 or bay 439006 or bay43 9006 or bay43-9006 or bay439006 or nexavar or 284461-73-0).mp,rn. (10800)
64 62 and 63 (5251)
65 exp kidney tumor/ (84645)
66 ((kidney or renal) adj3 (carcinoma$ or cancer$ or adenocarcinoma$ or neoplasm$ or tumo?r$ or malignan$ or metastas$)).mp. (93007)
67 ((nephroid or hypernephroid) adj3 carcinoma$).mp. (163)
68 (collecting duct adj3 carcinoma$).mp. (337)
69 (hypernephroma$ or RCC or mRCC or (tumor adj2 grawitz)).mp. (12459)
70 exp metastasis/ or metastas$.ab,ti. (414221)
71 exp Kidney/ or (renal or kidney).ab,ti. (844984)
72 70 and 71 (24013)
73 or/65-69,72 (110772)
74 64 and 73 (2786)
75 Random$.tw. or clinical trial$.mp. or exp health care quality/ (2895348)
76 animal/ (1797630)
77 animal experiment/ (1635745)
78 (rat or rats or mouse or mice or murine or rodent or rodents or hamster or hamsters or pig or pigs or porcine or rabbit or rabbits or animal or animals or dogs or dog or cats or cow or bovine or sheep or ovine or monkey or monkeys).mp. (5527603)
79 or/76-78 (5527603)
80 exp human/ (13853360)
81 human experiment/ (304202)
82 or/80-81 (13854795)
83 79 not (79 and 82) (4443855)
84 75 not 83 (2758037)
85 74 and 84 (1914)
86 85 not 61 (548)
87 limit 86 to embase (505)

Based on Trials filter (Best sens):
Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. Journal of the Medical Library Association 2006;94(1):41-7.

Cochrane Central Register of Controlled Trials (CENTRAL) (Cochrane Library Issue 8:2012) (Wiley)
Searched 3.9.12

#1 MeSH descriptor Carcinoma, Renal Cell explode all trees 399
#2 MeSH descriptor Kidney Neoplasms explode all trees 581
#3 (renal or kidney) NEAR/3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*) 1149
#4 hypernephroma* or nephroid carcinoma* or hypernephroid carcinoma* 8
#5 sunitinib or su?10398 or su ?10398 or sutent 115
#6 (#1 OR #2 OR #3 OR #4) 1166
#7 (#5 AND #6) 68
#8 progress* or fail* 74759

#9 interleukin 2 or bioleukin or IL-2 5277
#10 (alpha NEAR/2 interferon) or alferon or cilferon or kemron or veldona 4122
#11 cytokine* 5337
#12 (#5 OR #9 OR #10 OR #11) 12325
#13 (#8 AND #12) 2644
#14 (#5 OR #13) 2699
#15 (#6 AND #14) 211
#16 (#15), in 2012 10
#17 (#15 AND NOT #16) 201
#18 (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4) 119
#19 (sorafenib or bay 43 9006 or bay 43-9006 or bay 439006 or bay43 9006 or bay43-9006 or bay439006 or nexavar or 284461-73-0) 149
#20 (#18 AND #19) 32
#21 MeSH descriptor Carcinoma, Renal Cell, this term only 399
#22 MeSH descriptor Kidney Neoplasms, this term only 531
#23 ((kidney or renal) near/3 (carcinoma* or cancer* or adenocarcinoma* or neoplasm* or tumo?r* or malignan* or metastas*)) 1220
#24 ((nephroid or hypernephroid) near/3 carcinoma*) 1
#25 (collecting duct near/3 carcinoma*) 8
#26 (hypernephroma* or RCC or mRCC or (tumor near/2 grawitz)) 275
#27 MeSH descriptor Neoplasm Metastasis explode all trees 3305
#28 metastas* 7779
#29 (#27 OR #28) 7847
#30 MeSH descriptor Kidney explode all trees 2967
#31 (renal or kidney) 34823
#32 (#30 OR #31) 34829
#33 (#29 AND #32) 513
#34 (#21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #33) 1495
#35 (#20 AND #34) 29
#36 (#35 AND NOT #17) 2

CENTRAL search retrieved 0 records

Search B: (Sunitinib AND (Everolimus or Temsirolimus) AND mRCC AND RCTs) NOT (MS Search 10.15)

Medline search: Lines #1-52 replicate the MS search. The MS search was run in March 2012, so a date was inserted lines #53-55 in order to ensure any new papers retrieved by the original MS strategy would not be removed from the new results set. Lines #56-82 are the new ERG focused strategy intended to identify any RCTs containing Sunitinib AND (Everolimus OR Temsirolimus) AND mRCC. The MS search is then "NOT"-ed from the ERG search in lines 83 in order to remove papers already retrieved by the previous MS search, leaving only new or previously missed references.

Medline (OvidSP): 1946-2012/8/wk5
Searched 7.9.12

1 exp Carcinoma, Renal Cell/ (20040)
2 exp Kidney Neoplasms/ (54243)
3 ((renal or kidney) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).mp. (61216)
4 (sunitinib or su?10398 or su ?10398 or sutent).mp. (2158)

5 (hypernephroma* or nephroid carcinoma* or hypernephroid carcinoma*).mp. (1270)
6 Randomized controlled trials as Topic/ (82897)
7 Randomized controlled trial/ (336136)
8 Random allocation/ (75700)
9 Double blind method/ (116905)
10 Single blind method/ (16674)
11 Clinical trial/ (473453)
12 exp Clinical Trials as Topic/ (260609)
13 or/6-12 (837179)
14 (clinic$ adj trial$1).tw. (174415)
15 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw. (114527)
16 Placebos/ (31302)
17 Placebo$.tw. (138905)
18 Randomly allocated.tw. (13961)
19 (allocated adj2 random).tw. (675)
20 or/14-19 (355581)
21 13 or 20 (947185)
22 Case report.tw. (168907)
23 Letter/ (759065)
24 Historical article/ (286391)
25 Review of reported cases.pt. (0)
26 Review, multicase.pt. (0)
27 or/22-26 (1203839)
28 21 not 27 (920406)
29 Epidemiologic studies/ (5483)
30 exp case control studies/ (571339)
31 exp cohort studies/ (1207082)
32 Case control.tw. (62398)
33 (cohort adj (study or studies)).tw. (63494)
34 Cohort analy$.tw. (2821)
35 (Follow up adj (study or studies)).tw. (33433)
36 (observational adj (study or studies)).tw. (31985)
37 Longitudinal.tw. (112554)
38 Retrospective.tw. (218086)
39 Cross sectional.tw. (127163)
40 Cross-sectional studies/ (146450)
41 or/29-40 (1617289)
42 28 or 41 (2343805)
43 1 or 2 or 3 or 5 (63624)
44 4 and 42 and 43 (484)
45 (progress* or fail*).mp. (1388717)
46 (interleukin 2 or bioleukin or IL-2).mp. (67021)
47 ((alpha adj2 interferon) or alferon or cilferon or kemron or veldona).mp. (34666)
48 cytokine*.mp. (234597)
49 4 or 46 or 47 or 48 (307426)
50 45 and 49 (45380)
51 4 or 50 (46728)
52 42 and 43 and 51 (956)
53 2012$.ed. (561063)
54 52 and 53 (63)
55 52 not 54 (893)
56 (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4).af. (2189)

57 (temsirolimus or cci 779 or cci779 or nsc 683864 or nsc683864 or torisel or way-cci 779 or "cell cycle inhibitor 779" or 162635-04-3 or 343261-52-9).af. (737)
58 (everolimus or affinitor or afinitor or certican or "rad 001" or rad 001a or rad001 or rad001a or sdz rad or votubia or xience or zortress or 159351-69-6).af. (1919)
59 or/57-58 (2424)
60 56 and 59 (368)
61 carcinoma, renal cell/ or Kidney Neoplasms/ (51922)
62 ((kidney or renal) adj3 (carcinoma$ or cancer$ or adenocarcinoma$ or neoplasm$ or tumo?r$ or malignan$ or metastas$)).ti,ab,ot,hw. (62180)
63 ((nephroid or hypernephroid) adj3 carcinoma$).ti,ab,ot,hw. (112)
64 (collecting duct adj3 carcinoma$).ti,ab,ot,hw. (227)
65 (hypernephroma$ or RCC or mRCC or (tumor adj2 grawitz)).ti,ab,ot,hw. (8116)
66 exp neoplasm metastasis/ or metastas$.ab,ti. (263215)
67 exp kidney/ or (renal or kidney).ab,ti. (589117)
68 66 and 67 (12885)
69 or/61-65,68 (66542)
70 60 and 69 (275)
71 randomized controlled trial.pt. (336136)
72 controlled clinical trial.pt. (85111)
73 randomized.ab. (239114)
74 placebo.ab. (134375)
75 drug therapy.fs. (1566004)
76 randomly.ab. (171899)
77 trial.ab. (247764)
78 groups.ab. (1125860)
79 or/71-78 (2913768)
80 animals/ not (animals/ and humans/) (3686416)
81 79 not 80 (2475380)
82 70 and 81 (253)
83 82 not 55 (109)

Based on Trials filter:
Lefebvre C, Manheimer E, Glanville J. Chapter 6: searching for studies. Box 6.4.c: Cochrane Highly sensitive search strategy for identifying randomized controlled trials in Medline: Sensitivity-maximizing version (2008 version); OVID format. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org

Medline In-Process & Other Non-Indexed Citations (OvidSP): up to 2012/9/6
Medline Daily Update (OvidSP): up to 2012/9/6
Searched 7.9.12

1 (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4).af. (280)
2 (temsirolimus or cci 779 or cci779 or nsc 683864 or nsc683864 or torisel or way-cci 779 or "cell cycle inhibitor 779" or 162635-04-3 or 343261-52-9).af. (87)
3 (everolimus or affinitor or afinitor or certican or "rad 001" or rad 001a or rad001 or rad001a or sdz rad or votubia or xience or zortress or 159351-69-6).af. (244)
4 or/2-3 (292)
5 1 and 4 (63)
6 carcinoma, renal cell/ or Kidney Neoplasms/ (33)
7 ((kidney or renal) adj3 (carcinoma$ or cancer$ or adenocarcinoma$ or neoplasm$ or tumo?r$ or malignan$ or metastas$)).ti,ab,ot,hw. (2068)
8 ((nephroid or hypernephroid) adj3 carcinoma$).ti,ab,ot,hw. (0)

9 (collecting duct adj3 carcinoma$).ti,ab,ot,hw. (13)
10 (hypernephroma$ or RCC or mRCC or (tumor adj2 grawitz)).ti,ab,ot,hw. (627)
11 exp neoplasm metastasis/ or metastas$.ab,ti. (9011)
12 exp kidney/ or (renal or kidney).ab,ti. (16467)
13 11 and 12 (494)
14 or/6-10,13 (2338)
15 5 and 14 (48)
16 randomized controlled trial.pt. (689)
17 controlled clinical trial.pt. (45)
18 randomized.ab. (12711)
19 placebo.ab. (5008)
20 drug therapy.fs. (913)
21 randomly.ab. (12446)
22 trial.ab. (13394)
23 groups.ab. (72692)
24 or/16-23 (96255)
25 animals/ not (animals/ and humans/) (1423)
26 24 not 25 (95954)
27 15 and 26 (21)

Based on Trials filter:
Lefebvre C, Manheimer E, Glanville J. Chapter 6: searching for studies. Box 6.4.c: Cochrane Highly sensitive search strategy for identifying randomized controlled trials in Medline: Sensitivity-maximizing version (2008 version); OVID format. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org

Embase (OvidSP): 1974-2012/wk35
Searched 7.9.12

1 exp kidney carcinoma/ (39053)
2 exp kidney tumor/ (84645)
3 ((renal or kidney) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).mp. (90485)
4 (hypernephroma* or nephroid carcinoma* or hypernephroid carcinoma*).mp. (1508)
5 exp sunitinib/ (9234)
6 (sunitinib or su?10398 or su ?10398 or sutent).mp. (9504)
7 Clinical trial/ (875054)
8 Randomized controlled trial/ (330593)
9 Randomization/ (59294)
10 Single blind procedure/ (16335)
11 Double blind procedure/ (113064)
12 Crossover procedure/ (34836)
13 Placebo/ (216576)
14 Randomi?ed controlled trial$.tw. (78390)
15 Rct.tw. (9935)
16 Random allocation.tw. (1215)
17 Randomly allocated.tw. (17843)
18 Allocated randomly.tw. (1848)
19 (allocated adj2 random).tw. (792)
20 Single blind$.tw. (12773)
21 Double blind$.tw. (136713)
22 ((treble or triple) adj blind$).tw. (304)
23 Placebo$.tw. (184088)
24 Prospective study/ (212616)

25 or/7-24 (1290811)
26 Case study/ (16848)
27 Case report.tw. (239874)
28 Abstract report/ or letter/ (861436)
29 or/26-28 (1113428)
30 25 not 29 (1255043)
31 Clinical study/ (86264)
32 Case control study/ (70056)
33 Family study/ (9694)
34 Longitudinal study/ (55051)
35 Retrospective study/ (290790)
36 Prospective study/ (212616)
37 Randomized controlled trials/ (19708)
38 36 not 37 (212104)
39 Cohort analysis/ (129696)
40 (Cohort adj (study or studies)).mp. (87484)
41 (Case control adj (study or studies)).tw. (63467)
42 (follow up adj (study or studies)).tw. (42598)
43 (observational adj (study or studies)).tw. (46939)
44 (epidemiologic$ adj (study or studies)).tw. (68066)
45 (cross sectional adj (study or studies)).tw. (64024)
46 or/31-35,38-45 (1012257)
47 1 or 2 or 3 or 4 (102826)
48 30 or 46 (2001176)
49 5 or 6 (9504)
50 47 and 48 and 49 (2088)
51 (progress* or fail*).mp. (1866022)
52 (interleukin 2 or bioleukin or IL-2).mp. (97658)
53 ((alpha adj2 interferon) or alferon or cilferon or kemron or veldona).mp. (57445)
54 cytokine*.mp. (353364)
55 49 or 52 or 53 or 54 (460334)
56 51 and 55 (73843)
57 49 or 56 (79367)
58 47 and 48 and 57 (2835)
59 2012$.em. (879230)
60 58 and 59 (190)
61 58 not 60 (2645)
62 sunitinib/ or (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4).mp,rn. (9516)
63 temsirolimus/ or (temsirolimus or cci 779 or cci779 or nsc 683864 or nsc683864 or torisel or way-cci 779 or "cell cycle inhibitor 779" or 162635-04-3 or 343261-52-9).mp,rn. (4080)
64 Everolimus/ or (everolimus or affinitor or afinitor or certican or "rad 001" or rad 001a or rad001 or rad001a or sdz rad or votubia or xience or zortress or 159351-69-6).mp,rn. (8726)
65 63 or 64 (10741)
66 62 and 65 (2488)
67 exp kidney tumor/ (84645)
68 ((kidney or renal) adj3 (carcinoma$ or cancer$ or adenocarcinoma$ or neoplasm$ or tumo?r$ or malignan$ or metastas$)).mp. (93007)
69 ((nephroid or hypernephroid) adj3 carcinoma$).mp. (163)
70 (collecting duct adj3 carcinoma$).mp. (337)
71 (hypernephroma$ or RCC or mRCC or (tumor adj2 grawitz)).mp. (12459)
72 exp metastasis/ or metastas$.ab,ti. (414221)
73 exp Kidney/ or (renal or kidney).ab,ti. (844984)
74 72 and 73 (24013)

75 or/67-71,74 (110772)
76 66 and 75 (1496)
77 Random$.tw. or clinical trial$.mp. or exp health care quality/ (2895348)
78 animal/ (1797630)
79 animal experiment/ (1635745)
80 (rat or rats or mouse or mice or murine or rodent or rodents or hamster or hamsters or pig or pigs or porcine or rabbit or rabbits or animal or animals or dogs or dog or cats or cow or bovine or sheep or ovine or monkey or monkeys).mp. (5527603)
81 or/78-80 (5527603)
82 exp human/ (13853360)
83 human experiment/ (304202)
84 or/82-83 (13854795)
85 81 not (81 and 84) (4443855)
86 77 not 85 (2758037)
87 76 and 86 (1079)
88 87 not 61 (331)
89 limit 88 to embase (306)

Based on Trials filter (Best sens):
Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. Journal of the Medical Library Association 2006;94(1):41-7.

Cochrane Central Register of Controlled Trials (CENTRAL) (Cochrane Library Issue 8:2012) (Wiley)
Searched 7.9.12

#1 MeSH descriptor Carcinoma, Renal Cell explode all trees 399
#2 MeSH descriptor Kidney Neoplasms explode all trees 581
#3 (renal or kidney) NEAR/3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*) 1149
#4 hypernephroma* or nephroid carcinoma* or hypernephroid carcinoma* 8
#5 sunitinib or su?10398 or su ?10398 or sutent 115
#6 (#1 OR #2 OR #3 OR #4) 1166
#7 (#5 AND #6) 68
#8 progress* or fail* 74759
#9 interleukin 2 or bioleukin or IL-2 5277
#10 (alpha NEAR/2 interferon) or alferon or cilferon or kemron or veldona 4122 4122
#11 cytokine* 5337
#12 (#5 OR #9 OR #10 OR #11) 12325
#13 (#8 AND #12) 2644
#14 (#5 OR #13) 2699
#15 (#6 AND #14) 211
#16 (#15), in 2012 10
#17 (#15 AND NOT #16) 201
#18 (sunitinib or pha 2909040ad or pha2909040ad or "su 010398" or su010398 or "su 011248" or su011248 or su 10398 or su10398 or su 11248 or su11248 or sutent or 341031-54-7 or 557795-19-4) 119
#19 (temsirolimus or cci 779 or cci779 or nsc 683864 or nsc683864 or torisel or way-cci 779 or "cell cycle inhibitor 779" or 162635-04-3 or 343261-52-9) 56
#20 (everolimus or affinitor or afinitor or certican or "rad 001" or rad 001a or rad001 or rad001a or sdz rad or votubia or xience or zortress or 159351-69-6) 347
#21 (#19 OR #20) 399
#22 (#18 AND #21) 18
#23 MeSH descriptor Carcinoma, Renal Cell, this term only 399
#24 MeSH descriptor Kidney Neoplasms, this term only 531

#25 ((kidney or renal) near/3 (carcinoma* or cancer* or adenocarcinoma* or neoplasm* or tumo?r* or malignan* or metastas*)) 1220
#26 ((nephroid or hypernephroid) near/3 carcinoma*) 1
#27 (collecting duct near/3 carcinoma*) 8
#28 (hypernephroma* or RCC or mRCC or (tumor near/2 grawitz)) 275
#29 MeSH descriptor Neoplasm Metastasis explode all trees 3305
#30 metastas* 7779
#31 (#29 OR #30) 7847
#32 MeSH descriptor Kidney explode all trees 2967
#33 (renal or kidney) 34823
#34 (#32 OR #33) 34829
#35 (#31 AND #34) 513
#36 (#23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #35) 1495
#37 (#22 AND #36) 18
#38 (#37 AND NOT #17) 1

CENTRAL search retrieved 0 records

Appendix 1B: Critique of manufacturer's searches

Clinical effectiveness

Limitations

- The ERG queried the following missing details for the Web of Science search with the MS and received the following responses to the clarification letter:
    - Host = DataStarWeb
    - Indices searched= Science Citation Index
    - Date range searched=1900-2010/07/08
    - Search dates= 8th July 2010
- The ERG noted some inconsistency regarding the use of parentheses in the Web of Science search. Parentheses are present in line #72, but missing in subsequent lines i.e. #79 and #83 etc. The ERG has no access to this host, so was unable to ascertain if this would have been negated by the order of precedence in which the host would have searched for the Boolean operators, so cannot confirm if this would impacted on results
- The ERG noted CAS registry numbers were not included in the search strategies
- The ERG noted the absence of the following brand names:
    - Pazopanib missing: votrient
    - Interferon alpha missing: varients using alfa rather than alpha
    - Everolimus missing: affinitor or afinitor or xience or zortress
    - Axitinib missing: Inlyta
- There was some redundant usage of the explosion function in both the Embase and Medline searchs (see Embase lines #5 & #7 for examples)
- No host was reported for the Cochrane Library searches, but given the syntax used the ERG presumed this to be Wiley.

Additional Search (10.15)

Limitations

- The ERG noted that in places the logic behind the searches appears unclear (10.15). All strategies contained an orphan line (Line #44 Medline strategy), which appeared to be made redundant by the final set. Each search also appeared to contain a redundant use of the drug Sunitinib in a line combined using OR with cytokines (see line #49 in Medline), which was superseded by inclusion of the drug without qualifiers in the penultimate line (see #51 Medline). However these errors did not appear to be consequential to the final results.
- There appeared to be a typographical error for date reported for the Medline search in 10.15.2. searches were reported as being performed on 14th March 2012, whilst 10.15.1 reported 14th March 2011, given the rest of the searches in this section were documented as being run in March 2012, the ERG assumed this to be the correct date.
- As with Clinical Effectiveness searches, the ERG noted the absence of CAS registry numbers and missing variant spellings for Interferon alpha, using *alfa*.

Indirect and mixed treatment comparisons

The MS reported that the strategies presented in 10.2 were employed for this section. The same limitations applied.

Adverse events (comparators)

In utilising the same strategies reported in 10.2 the same limitations applied. Given the CRD advice on not using RCT filters in these cases, the ERG would recommend removing the study designs filters in lines #17-93 (Medline Search) and replacing them with a suitable adverse events filter, a number of which can be found in the ISSG Search Filters Resource.[49]

Non-RCT Evidence (Apixaban)

Adequate searches were carried out on all NICE required databases. ERG noted the same limitations in the line for Axitinib as in earlier searches (see Clinical Effectiveness 10.2). A test by the ERG in Medline and Embase showed that the omission of the alternative name Inlyta was unlikely to have impacted on recall.

	# ▲	Searches	Results	Search Type	Actions
☐	1	(axitinib or ag13736 or ag 13736).mp.	160	Advanced	Display More »
☐	2	(axitinib or ag13736 or ag 13736 or Inlyta).mp.	160	Advanced	Display More »

Remove Selected | Save Selected | Combine selections with: And Or RSS

Save Search History

Cost effectiveness

Limitations

- The ERG noted the redundant use of economics filter on Cochrane searches
- The ERG noted some disparity between the way that renal cell carcinoma was searched for between this and the earlier clinical effectiveness searches (10.2), especially the omission of terms such as hypernephroma$ or nephroid carcinoma$. It was not clear if this had an impact on the overall recall of the strategies
- The ERG noted that the following Emtree and MeSH terms used in the previous clinical effectiveness searches were not used in the cost effectiveness searches: Emtree *interleukin 2/* (Embase) and the MeSH for both *Interferon-Alpha/* and *Interleukin-2/* (Medline and Cochrane). It was unclear if this had any impact on the recall of results.
- The ERG noted the absence of CAS registry numbers.
- The ERG noted the absence of the following brand names:
    - Tivozanib missing: krn951
    - Pazopanib missing: votrient
    - Interferon alpha missing: varients using alfa rather than alpha
    - Everolimus missing: affinitor or xience or zortress
- No host was reported for the Embase and Econlit searches, ERG assumed this to be Ovid as reported for Medline search
- No host was given for the Cochrane Library searches, but given the syntax used and after a brief investigation of the hits per line, the ERG assumed this to be EBM Reviews from Ovid.
- Results were limited to English language only, which may have resulted in the omission of potentially useful papers in other languages.

Measurement and valuation of health effects

Limitations

- Redundant use of HRQL filter on Cochrane searches

- The ERG noted some disparity between the way that renal cell carcinoma was searched for between this and the earlier clinical effectiveness searches, especially the omission terms such as hypernephroma$ or nephroid carcinoma$. It was not clear if this would have impacted the overall recall of the strategies
- Search Strategies in this section also appeared to be missing the HRQoL outcome SF-6D. These may have been retrieved by the HRQL facet in the earlier Clinical Effectiveness searches (10.2), which did include this outcome
- No host was reported for the Embase and Econlit searches, ERG assumed this to be Ovid as reported for Medline search
- No host was given for the Cochrane Library searches, but given the syntax used and after a brief investigation of the hits per line, the ERG assumed this to be EBM Reviews from Ovid.
- Results were limited to English language only, which may have resulted in the omission of potentially useful papers in other languages.

Resource identification, measurement and valuation

Please see section 4.1.1.

Appendix 2: Summary list of cost-effectiveness evaluations

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Aiello et al[50]	2007	Argentina	To estimate the cost-effectiveness of sunitinib malate versus BSC in the treatment of cytokine-refractory mRCC patients	Cytokine-refractory metastatic RCC patients failing on IL-2, IFN-alpha or combination of these.	Markov model. Effectiveness results and utility data were taken from a clinical trial and a US Medicare database. Data was adjusted with general population mortality estimates from Argentinean life tables.	Discounted: 0.98 QALY (sunitinib vs BSC)	Discounted: AR$52,243 (sunitinib vs BSC)	AR$53,445 per QALY (sunitinib vs BSC)
Casciano et al[51]	2011	USA	To examine the potential cost-effectiveness of everolimus vs sorafenib therapy for the treatment of metastatic renal cell carcinoma after failure of first-line sunitinib from a US payer perspective	Patients with metastatic RCC after failure of first-line sunitinib.	Markov model. Time horizon of 6 years with 8-week cycles. Four health states: SD no AEs, SD with AEs, PD, death. Transition probabilities based on analysis of patient-level data from RCT and single-arm trial, utilities from the PenTAG (UK analysis) report.	Discounted: 0.916 QALY (everolimus vs sorafenib)	Discounted: $81,643 (everolimus vs sorafenib)	$89,160 per QALY (everolimus vs sorafenib)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Contreras-Hernandez et al[52]	2007	Mexico	To compare the economic and health consequences of sunitinib vs BSC in adult patients with mRCC who failed prior cytokine treatment from a health care payer's perspective in Mexico	Adult patients failing cytokine therapies with metastatic RCC in stages III and IV.	Markov model. Time horizon of ten years. Four health states: no new progression, death due to metastatic RCC, history of new progression, death due to other causes. Transition probabilities and QALYs obtained according to clinical trials from the published literature.	Discounted: sunitinib: 1.32 QALYs; BSC: 0.39 QALYs	Discounted: sunitinib: US$36,928; BSC: US$4,103	US$35,238 per QALY (sunitinib vs BSC)
El Ouagari et al[53]	2010	Canada†	To compare the cost-effectiveness of everolimus vs BSC in patients who failed on VEGF-TKI therapy from a Canadian societal perspective	Metastatic RCC patients whose disease failed on VEGF-TKI therapies.	Markov model simulating 2 hypothetical patient cohorts, using a 6 year time horizon. Health state transition probabilities were derived from a RCT and costs and utilities were drawn from literature.	Discounted: 0.469 QALY (everolimus vs BSC)	Discounted: $29,080 (everolimus vs BSC)	$62,067 per QALY (everolimus vs BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Gao et al[54]	2006	USA	To evaluate the cost –effectiveness of sorafenib + BSC versus BSC alone in advanced RCC from a UK payer perspective	Patients with advanced RCC‡.	Markov model to project lifetime survival. Three health states: PFS, progression, death. Transition probabilities were obtained from a RCT.	Not reported	Lifetime per patient, discounted: sorafenib + BSC: $85,571; BSC: $36,634	Not reported $75,354 per LYG (sorafenib + BSC vs BSC)
Gao et al[§55]	2008	USA	To update the earlier economic model (reported in [56] with the latest clinical data to evaluate the cost-effectiveness of sorafenib + BSC versus BSC alone in advanced RCC from a US payer perspective	Patients with advanced RCC‡.	Markov model to project lifetime survival. Three health states: PFS, progression, death. Transition probabilities were obtained from a RCT.	Not reported 0.88 discounted life years (sorafenib + BSC vs BSC)	Lifetime per patient, discounted: sorafenib + BSC: $92,222; BSC: $36,634	Not reported $63,219 per LYG (sorafenib + BSC vs BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Hoyle et al[33]	2010	UK	To estimate the cost-effectiveness of sorafenib vs BSC for the second-line treatment of advanced renal cell carcinoma from the perspective of the UK NHS	Patients with advanced RCC, resistant to standard therapy; 82% had previously received cytokine-based therapy.	Markov model with a 10-year time horizon and 6-week cycles. Three health states: PFS, PD, death. Utilities were derived from a phase II single-arm trial of sunitinib. Clinical effectiveness from a RCT of sorafenib vs placebo.	Discounted: 0.27 QALY (sorafenib vs BSC)	Discounted: £20,063 (sorafenib vs BSC)	£75,398 per QALY (sorafenib vs BSC)
Jaszewski et al[57]	2007	Canada	To evaluate the cost-effectiveness of sorafenib + BSC vs BSC alone in advanced RCC from a Canadian provincial Ministry of Health perspective	Patients with advanced RCC‡.	Markov model to project lifetime survival. Three health states: PFS, progression, death.	Not reported	Lifetime per patient, discounted: sorafenib + BSC: CAD$62,426; BSC: CAD$18,898	Not reported CAD$36,046 per LYG (sorafenib + BSC vs BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Ondrackova et al[¶58]	2010	Czech Republic	To assess the cost-effectiveness of sorafenib and sunitinib for the treatment of mRCC in reimbursement proceedings vs data from clinical practice	Patients with advanced or metastatic RCC after cytokine intolerance or failure.	Not reported The study compared cost-effectiveness results from manufacturers' submissions with own analysis results based on patient data from comprehensive cancer centre clinical practice (comparator: 70% treated with sunitinib and 30% treated with BSC).	Not reported	Not reported	Not reported Manufacturer submission: €37,143 per progression-free year (sorafenib vs sunitinib or BSC) New analysis: €19,878 per progression-free year (sorafenib vs sunitinib or BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Paz-Ares et al[59]	2010	Spain	To investigate the cost-effectiveness of sunitinib vs BSC in patients with cytokine refractory mRCC from the perspective of the Spanish NHS	Patients with metastatic RCC who did not respond to, were intolerant to or experienced disease progression on IL-2 or IFN-alpha. The model included characteristics of a Spanish population: average age of 62 years and 66% men and 34% women.	Markov model with a 10-year time horizon and a 4-week cycle length. Three health states: PFS, survival with progression, death from metastatic RCC or other causes. Utilities and effectiveness data were obtained from a phase II study of sunitinib-treated patients.	Discounted: sunitinib: 1.36 QALYs; BSC: 0.39 QALYs	Discounted: €32,911 (sunitinib vs BSC)	€34,196 per QALY (sunitinib vs BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Pitt et al[††60]	2010	UK	ERG report on the manufacturer's submission: Single Technology Appraisal for everolimus (Afinitor®) in advanced renal cell carcinoma	Heavily pre-treated adult (≥18 years) advanced RCC patients who have experienced disease progression on or following one or more VEGF-targeted therapies (sunitinib, sorafenib and/or bevacizumab).	Markov model with 8-week cycles and a 144 week-time horizon. Four health states: stable disease with AEs, stable disease without AEs, progressed, death. Utility data from the PenTAG report was used and effectiveness data was obtained from a RCT.	Manufacturer submission: With patient-access scheme, discounted: 0.304 QALY (everolimus + BSC vs BSC) ERG re-analysis: With patient-access scheme, discounted: 0.193 QALY (everolimus + BSC vs BSC)	Manufacturer submission: With patient-access scheme, discounted: £15,704 (everolimus + BSC vs BSC) ERG re-analysis: With patient-access scheme, discounted: £12,610 (everolimus + BSC vs BSC)	Manufacturer submission: With PAS: £51,613 per QALY (everolimus + BSC vs BSC) ERG re-analysis: With PAS: £65,231 per QALY (everolimus + BSC vs BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Purmonen et al[61]	2008	Finland	To analyse the cost-effectiveness of sunitinib as second-line therapy for cytokine-refractory mRCC compared with BSC in Finland	Patients with metastatic RCC (median age 68 years), previously treated with IFN- (69% male, 31% female).	Markov model with 5-year time horizon and 1-month cycles. Three health states: no new progression events, history of progression-related events, death. Transition probabilities and utilities were obtained from a phase II single-arm trial and a beta distribution was used for uncertainty regarding BSC utilities.	Discounted: 0.74 QALY (sunitinib vs BSC)	Discounted, per patient, for 5 years: €32,630 (sunitinib vs BSC)	€43,698 per QALY (sunitinib vs BSC)
Tatar et al[62]	2009	Turkey	To assess the cost-effectiveness of sorafenib + BSC versus BSC alone in mRCC patients in Turkey	Patients with unresectable and/or metastatic RCC‡.	Markov model over a patient's lifetime. Three health states: PFS, disease progression, death. PFS and survival were extrapolated from a RCT.	Not reported 1.269 discounted LYG (sorafenib + BSC vs BSC)	Lifetime per patient, discounted: sorafenib + BSC: 47,665 TL; BSC: 4,080 TL	Not reported 34,342 TL per LYG (sorafenib + BSC vs BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Teich et al[63]	2009	Brazil	To develop a cost-effectiveness analysis of sorafenib + BSC vs BSC alone in the second-line treatment of advanced RCC from the Brazilian public health care system perspective	Advanced RCC (second-line treatment).	Markov model with a lifetime time horizon and a 3-month cycle. Three health states: PFS, disease progression, death. Transition probabilities were obtained from a RCT.	Not reported Mean PFS: 2066 years (sorafenib/BSC); 1243 years (BSC)	Lifetime, discounted: sorafenib/BSC: R$48,285; BSC: R$7,356	Not reported R$49,751 (US$21,553) per LYG (sorafenib/BSC vs BSC)
Thompson-Coon et al[§§64]	2010	UK	To assess the cost-effectiveness of bevacizumab, combined with IFN, sorafenib tosylate, sunitinib and temsirolimus in the treatment of advanced/mRCC	All patients in the model were assumed to have advanced/metastatic RCC and all patients were assumed to start in PFS.	Markov model with 10-year time horizon and 6-week cycles. Three health states: progression-free, progressive, death. Weibull curves were fitted to empirical effectiveness data from a RCT. Utility data was obtained from manufacturer submissions.	Discounted: 0.23 QALY (sorafenib vs BSC)	Discounted: £24,001 QALY (sorafenib vs BSC)	£102,498 per QALY (sorafenib vs BSC)

Study	Year	Country where study was performed	Aim	Patient population (average age in years)	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Van Nooten et al[65]	2007	Belgium	To determine the cost-effectiveness of sunitinib malate vs BSC after failure of cytokine immunotherapy from the perspective of the Belgian public payers	Patients with metastatic RCC after failure on first-line cytokine therapy.	Markov model with a 10-year time horizon and a one-month cycle length. Three health states: PFS, tumour progression and move to BSC, death. Effectiveness parameters for sunitinib were taken from a phase II clinical trial. Utilities were derived from published literature.	Not reported Average discounted 1.11 LYG per patient (sunitinib vs BSC)	Not reported	Not reported €35,389 per LYG (sunitinib vs BSC)

Abbreviations: AE, adverse event; BSC, best supportive care; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; IFN, interferon; IL, interleukin; LYG, life-year gained; mRCC, metastatic renal cell carcinoma; NHS, National Health Service; PAS, patient access scheme; PenTAG, Peninsula Technology Assessment Group; PD, progressive disease; PFS, progression-free survival; QALY(s), quality-adjusted life year(s); RCC, renal cell carcinoma; RCT, randomised controlled trial; SD, stable disease; TKI, tyrosine kinase inhibitor; VEGF, vascular endothelial growth factor.

†It is not clear from this abstract if the costs are in US$ or CA$.

‡It is not clear from the abstract if the patients are second-line treated, but it was assumed that the patients were second-line RCC patients, because data was used from the (second-line) TARGET trial.

§This analysis updated the Gao et al 2006 analysis, using latest overall survival data.

¶This abstract describes two cost-effectiveness analyses: first-line sunitinib vs interferon-alpha and second-line sorafenib vs sunitinib or BSC. Only the second-line data is included in the systematic review.

††This ERG report was identified in the electronic database searches and the relevant manufacturer submission to NICE for everolimus was consulted for additional information.

§§This HTA document reports various comparisons, however only the second-line treatment comparison (sorafenib vs BSC) is included in the systematic review. The manufacturer submissions and the ERG report (PenTAG report) were also consulted for additional information

Appendix 3: Phillips et al. Checklist

Results of assessing the manufacturers report based on the checklist by Phillips et al.

1. **Is there a clear statement of the decision problem?**

   Yes, the decision problem is clearly stated. (see section 5 of MS)

2. **Is the objective of the evaluation and model specified consistent with the stated decision problem?**

   Yes.

3. **Is the primary decision-maker specified?**

   The term is not used, but implicitly the NHS is assumed.

4. **Is the perspective of the model stated clearly?**

   Yes, the NHS perspective.

5. **Are the model inputs consistent with the stated perspective?**

   Yes, though the valuation of the EQ-5D descriptives are based on a US tariff.

6. **Has the scope of the model been stated and justified?**

   Yes.

7. **Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?**

   Yes.

8. **Is the structure of the model consistent with a coherent theory of the health condition under evaluation?**

   Yes. The structure of the model is based on previously identified models of advanced/mRCC and was validated by UK clinical expert opinion.

9. **Are the sources of data used to develop the structure of the model specified?**

   Yes.

10. **Are the causal relationships described by the model structure justified appropriately?**

    Yes.

11. **Are the structural assumptions transparent and justified?**

    Yes.

12. **Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?**

    Yes.

13. **Is there a clear definition of the options under evaluation?**

    Yes.

14. **Have all feasible and practical options been evaluated?**

    Yes.

15. **Is there justification for the exclusion of feasible options?**

    NA.

16. **Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?**

    Yes.

17. Is the time horizon of the model sufficient to reflect all important differences between options?

Yes. A lifetime horizon of 10 years was chosen in keeping with previous technology appraisals in mRCC.

18. Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?

Yes.

19. Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?

Yes.

20. Is the cycle length defined and justified in terms of the natural history of disease?

Yes.

21. Are the data identification methods transparent and appropriate given the objectives of the model?

Yes, with the exception of the method described in the MS to compute the mean progression-free utility, this could not be replicated.

22. Where choices have been made between data sources, are these justified appropriately?

Yes.

23. Has particular attention been paid to identifying data for the important parameters in the model?

Yes.

24. Has the quality of the data been assessed appropriately?

Yes.

25. Where expert opinion has been used, are the methods described and justified?

Not completely. According to the MS (Section 7.3.5), expert opinion was solicited (from one clinical expert and one health economic expert) to test and verify the choice of extrapolation method for OS and PFS curves; the methodology and results of the indirect comparison and STC approaches; the resource utilisation estimates for routine medical management and management of adverse events; and the utility estimates. In addition, further input was sought during an advisory board with five UK clinicians to validate the Axitinib model arm extrapolations, as well as the STC and RENCOMP comparisons. However, it is not mentioned in the MS how expert opinion has been elicited and incorporated in the study.

26. Is the data modelling methodology based on justifiable statistical and epidemiological techniques?

Yes, except for the simulated treatment comparison (STC). As mentioned in Section 1.5 of this report, whether an STC presents a valid and reliable estimate of the clinical effectiveness of axitinib versus BSC in a sunitinib refractory population is hard to determine. As there is no direct trial evidence it is not possible to compare the results of the STC to any existing evidence so the accuracy and reliability of the results cannot be ascertained.
Especially problematic is the fact that no measures of uncertainty are presented with the results of the STC

27. Is the choice of baseline data described and justified?

Yes.

28. Are transition probabilities calculated appropriately?

Yes.

29. Has a half-cycle correction been applied to both cost and outcome?

Yes.

30. If not, has this omission been justified?

NA.

31. If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?

Yes for the cytokine refractory subgroup. For the sunitinib subgroup, see 26.

32. Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?

Yes.

33. Have alternative extrapolation assumptions been explored through sensitivity analysis?

Yes.

34. Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?

Partly, most patients stop treatment when progression occurs. In that case, survival of patients is still modelled through the OS. However, the model assumes that patients withdrawing from treatment due to side effects follow the same PFS as patients on treatment. This is only valid if these patients continue to be followed up for progression. Whether this is indeed the case cannot be stated with certainty.

35. Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?

No, as patients withdraw steadily over time, the current model structure did not allow for such analysis.

36. Are the costs incorporated into the model justified?

Yes.

37. Has the source for all costs been described?

Yes.

38. Have discount rates been described and justified given the target decision-maker?

Yes.

39. Are the utilities incorporated into the model appropriate?

Yes.

40. Is the source for the utility weights referenced?

The utility weights are based on the EQ-5D administered to patients in the AXIS trial which were then values using a US tariff. The valuation part is not described in the MS but it is in the clinical study report of the AXIS trial.

41. Are the methods of derivation for the utility weights justified?

Yes, though the use of US values versus UK values may have introduced a small bias.

42. Have all data incorporated into the model been described and referenced in sufficient detail?

No. The method described in the MS to reproduce the mean progression-free utilities is unclear. The costs of death are described in the clarification letter (Section B – Question 21) but not in the MS. For some parameters for the probability distributions, standard deviations instead of standard errors were used.

43. Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?

NA

44. Is the process of data incorporation transparent?

Yes.

45. If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?

Yes

46. If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?

No, several error were made where standard deviations were interpreted as standard errors

47. Have the four principal types of uncertainty been addressed?

No. Methodological uncertainty is not discussed but a sensitivity analysis was done on discount rates. Structural uncertainty was explored through different scenario analyses. No assessments of heterogeneity were performed. Parameter uncertainty has been studied in the PSA.

48. If not, has the omission of particular forms of uncertainty been justified?

Yes. The MS Section 6.7.9 states that "no assessments of heterogeneity were performed as there was only a single study available for each pair-wise comparison. A network-meta analysis could not be performed due to a lack of trials that linked between different treatments and therefore no testing of inconsistency was possible".

49. Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?

No.

50. Is there evidence that structural uncertainties have been addressed via sensitivity analysis?

Yes. Several alternative scenarios have been run for different structural assumptions in order to explore their impact on the model outcomes.

51. Has heterogeneity been dealt with by running the model separately for different subgroups?

Yes, with the model, analyses are done for the cytokine refractory and sunitinib refractory subgroup separately.

52. Are the methods of assessment of parameter uncertainty appropriate?

Yes.

53. If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?

No. They were clearly stated but not justified. Moreover, in the MS parameter values were varied ±20% to the base case value. In response to the clarification letter (Section B – Question 25), the univariate sensitivity analysis was performed with parameter variation based on 95% confidence intervals.

54. Is there evidence that the mathematical logic of the model has been tested thoroughly before use?

No, in the MS it is stated that a rigorous quality check of the model was performed by a peer-reviewer not involved in the model development. The spread sheet provided which should have included the specific tasks performed, and their results returned only contained a post-testing list of bugs found in the model and the corresponding action to fix them.

55. Are any counterintuitive results from the model explained and justified?

No. There is one counterintuitive result and that is that the uncertainty around the QALY gain for the cytokine refractory subgroup is much larger than for the sunitinib refractory subgroup, despite the fact that more data was available to inform the first assessment. This is neither observed nor explained in the MS.

56. If the model has been calibrated against independent data, have any differences been explained and justified?

NA

57. Have the results of the model been compared with those of previous models and any differences in results explained?

No, the OS outcomes in BSC were compared to those from other studies in BSC, but the source of differences was not explored. No comparisons were made for the PFS outcomes in BSC.