

## **CONFIDENTIAL UNTIL PUBLISHED**

### **Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE**

### **Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma**

#### **ERRATUM**

Replacement pages following the factual accuracy check by  
Ipsen

<b>Produced by</b>	Southampton Health Technology Assessments Centre (SHTAC)
<b>Authors</b>	Dr Jonathan Shepherd, Principal Research Fellow Professor Joanne Lord, Professorial Fellow in Health Economics. Ms Petra Harris, Research Fellow Mr Olu Onyimadu, Research Fellow (Health Economics) Dr Geoff Frampton, Senior Research Fellow Maria Chorooglou, Senior Research Fellow (Health Economics)
<b>Correspondence to</b>	Dr Jonathan Shepherd Southampton Health Technology Assessments Centre (SHTAC) Wessex Institute Alpha House Enterprise Road, University of Southampton Science Park Southampton SO16 7NS <a href="http://www.southampton.ac.uk/shtac">www.southampton.ac.uk/shtac</a>

## **SUMMARY**

### **Scope of the company submission**

The company submission (CS) presents evidence for the clinical effectiveness and cost effectiveness of cabozantinib (CABOMETYX®) for the first-line treatment of patients with untreated locally advanced or metastatic renal cell carcinoma (RCC). Cabozantinib is an orally administered tyrosine kinase (RTK) inhibitor. The drug inhibits vascular endothelial growth factor (VEGF) and hepatocyte growth factor receptor protein (MET), implicated in tumour growth and angiogenesis, pathologic bone remodelling, drug resistance, and metastatic progression of cancer. The recommended dose is 60 mg once daily, with lower dose adjustments recommended to manage adverse reactions. Treatment continues until disease progression or the occurrence of unacceptable toxicity.

The patient population in the CS is adults with untreated, intermediate or poor risk (International Metastatic RCC Database Consortium (IMDC) criteria), locally advanced or metastatic RCC. The CS reports a comparison of the effects of cabozantinib versus sunitinib and versus pazopanib as initial therapy for patients with poor or intermediate risk metastatic RCC.

### **Summary of submitted clinical effectiveness evidence**

Systematic literature searches were performed to identify relevant clinical effectiveness studies. Searches identified one randomised controlled trial (RCT) of relevance to the appraisal, the CABOSUN trial. No direct trial evidence comparing cabozantinib versus pazopanib was identified.

CABOSUN was an investigator-led open-label, phase II RCT conducted by the Alliance for Clinical Trials in Oncology and conducted in 77 centres in the USA. It compared cabozantinib against sunitinib as first-line treatment. The trial included adult patients ( $\geq 18$  years of age) with untreated clear cell metastatic RCC, Eastern Cooperative Oncology Group (ECOG) performance status of 0 to 2 and intermediate or poor risk per IMDC criteria. Patients received 60 mg of cabozantinib (n=79) orally once per day or 50 mg of sunitinib (n=78) orally once per day (sunitinib: 4 weeks on and 2 weeks off), with treatment cycles for both trial arms defined as 6 weeks. Although not designed as a registration trial, the trial was used to support the marketing authorisation for cabozantinib for this indication (anticipated date of approval: May 2018) based on what the CS describes as “encouraging findings”. The trial is a key source of evidence for the company’s cost-effectiveness analysis. Based on the

requirements for the marketing authorisation, the CS presents retrospective analysis of this trial using assessment of tumour response and progression by an independent radiology committee (IRC), and using US Food and Drug Administration (FDA)-recommended censoring rules.

The primary trial outcome measure was progression free survival (PFS). Secondary outcome measures included: overall survival (OS), objective response rate (ORR) and adverse effects (AE) of treatment. Patient cross-over between trial arms was not permitted during the trial, however, upon disease progression patients in both arms received subsequent systemic non-radiation anti-cancer treatments (cabozantinib group 57%; sunitinib group 58%). Health-related quality of life (HRQoL) was not measured in the trial (alternative sources of HRQoL utility estimates were used in the economic model).

Generally, baseline characteristics between the treatment arms were balanced apart from the proportion of patients with  $\geq 2$  metastatic sites (cabozantinib group 79%; sunitinib group 67%).

Outcome data from the CABOSUN trial were reported for different data cut-off points. The ERG presents data in this report for the latest time-point available for each outcome: PFS - September 2016; OS - January 2017 and an updated analysis July 2017; and tumour response - September 2016.

#### Results of the CABOSUN trial

##### PFS

- At a median follow-up of 25 months (September 2016 data cut-off), median PFS was 8.6 months (95% confidence interval (CI) 6.8, 14.0) for cabozantinib and 5.3 months (95% CI 3.0, 8.2) for sunitinib ( $p=0.0008$ ), with a median difference of 3.3 months.
- The hazard ratio (HR), stratified by IMDC risk category and bone metastases, was 0.48 (95% CI 0.31, 0.74).
- The majority of events recorded were for documented disease progression (cabozantinib 51%, sunitinib 55%). PFS at 12 months (% event free) was 43.1 and 21.1 in the cabozantinib and sunitinib groups, respectively.

##### OS

At a median follow-up of 28.9 months for OS (January 2017 data cut-off), the median OS was 30.3 months (95% CI 14.6, not estimable) in the cabozantinib arm versus



- for why the curves should come together and then diverge between about 13 and 20 months.
- Median survival for OS and hazard ratio estimates are less favourable for the most recent data cut-off (July 2017) than in the earlier cut-off of January 2017 used to fit OS in the model (CS B.2.6 Figures 6 and 7). (NB. The CS states that the OS January 2017 dataset was used to inform in the model, and the KM plot is reproduced in the CS economic chapter). This suggests that the model may over-estimate the survival advantage for cabozantinib over sunitinib.
- The ERG considers that it is highly unlikely that the QALY loss is the same for all types of TEAE, but that these assumptions reflect a reasonable average. We conduct additional scenario analysis to test model sensitivity to the TEAE disutility parameter, including higher as well as lower estimates of the disutility. In addition, we note that of 59 types of adverse events listed in the company's model, only 18 events with incidences equal to or greater than 5% were modelled. We test the impact of changing the inclusion threshold for TEAEs in scenario analysis.
- The model does not include an adjustment for age-related increase in mortality in the general population, as the model relies entirely on the projected OS curves. However, given the high rate of mortality for people with advanced RCC, this might not affect results. We check that the model does not yield counter-intuitive results with longer-surviving RCC patients having lower mortality than members of the general population at the same age.

### **Summary of additional work undertaken by the ERG**

We corrected the company's model to reflect the identified errors. The most significant were coding errors in QALY calculations that had the effect of underestimating QALYs for each treatment, and hence underestimating the incremental QALY gain with cabozantinib compared with sunitinib and pazopanib. There were also small discrepancies in some cost estimates. The corrected model resulted in lower ICER estimates for the company's base case:

- £31,956 per QALY for the direct comparison of cabozantinib with sunitinib;
- £40,757 for cabozantinib compared with pazopanib and £26,182 compared with sunitinib based on the ITC analysis.

These estimates are subject to uncertainty, with the method of fitting the OS curves and choice of survival function having the largest impact on the ICERs.

Probabilistic analysis estimated a 28% probability of the ICER compared with pazopanib being less than £30,000 per QALY gained in the ITC base case.

We conducted additional analyses to test alternative assumptions and scenarios. The ERG-preferred set of assumptions included the following key differences from the company base cases:

- **Method of fitting OS curves.** Due to our concerns about the robustness of the ITC, we prefer to rely on the analysis of CABOSUN data for direct comparison of cabozantinib with sunitinib. Although the proportional hazards assumption appears not to hold, we agree with the company that the exponential distribution gives the best balance of fit to the trial data for both treatment arms and plausible long-term extrapolations. We base the OS curve for sunitinib on the exponential curve fitted to CABOSUN data. We then estimate the cabozantinib OS curve using the reported hazard ratio from the most recent update of trial data (July 2017 data cut) – the company’s analysis uses an earlier dataset (January 2017). Finally, based on the relationship shown in COMPARZ we assume equivalent OS for pazopanib and sunitinib, using the evidence from CABOSUN.
- **PFS and TTD curves.** We follow the company’s direct base case for estimates of PFS and TTD for cabozantinib and sunitinib: with lognormal curves separately fitted by treatment to CABOSUN data. For pazopanib, we again assumed equivalence with sunitinib for time to progression based on the results of the COMPARZ trial.
- **Time horizon and duration of effects.** The company uses a 20 year time horizon, which is longer than in other recent appraisals for RCC. We believe that it is correct to reflect a whole life time horizon, so also use 20 years in our base case. However, we do not believe that it is appropriate to assume persistence of treatment effects for cabozantinib based on the limited trial follow-up and sample size. The ERG therefore adopts a conservative assumption that progression and mortality hazards for cabozantinib equal those of sunitinib after a fixed period of time: 5 years from baseline in our preferred analysis.
- **Health state utilities, adverse effects and costs.** The company approach to modelling the utility and cost impacts of the treatments were generally reasonable and reflected the NICE base case and decisions in previous appraisals. We therefore adopt the same base case parameters, but conduct some additional scenario analyses to test the robustness of the results.

The ERG preferred analysis gave estimated ICERs of £65,742 for cabozantinib compared with pazopanib and £41,465 compared with sunitinib (Table 3). As in the company base

- Favourable – 0 factors
- Intermediate - 1 or 2 factors
- Poor -  $\geq 3$  factors.

The IMDC model has been externally validated in patients with metastatic RCC who were treated with first-line VEGF-targeted treatment, including patient stratification by risk (favourable risk group median overall survival 43.2 months after the start of targeted treatment, intermediate risk group 22.5 months and poor risk group 7.8 months).<sup>7</sup> The CS states that around 80% of all metastatic RCC patients are in the latter two risk groups and clinical experts advising the ERG concur with this. The CS cites a 5-year relative survival rate for stage IV RCC (i.e. metastatic) by Cancer Research UK as around 6% in the UK.<sup>2</sup>

### **2.1.3 Effects of RCC on health-related quality of life**

The top five symptoms reported in a national, cross-sectional study by patients with advanced metastatic RCC are: fatigue, weakness, worry, shortness of breath, and irritability.<sup>9</sup> HRQoL in this patient group is also impaired by disease-related factors associated with tumour burden, for example anorexia-cachexia syndrome (associated with weight loss, lethargy, as well as possible fever, night sweats and distortion of the sense of taste amongst others), hypercalcemia, venous thromboembolism, pain (somatic, visceral and neuropathic), and metastases-associated specific site symptoms.<sup>10</sup>

Patients with advanced RCC generally have a poor prognosis and this, combined with the symptoms associated with advanced disease, can significantly affect all domains of patients' HRQoL not just physical functioning, such as emotional and social wellbeing and.<sup>10 11</sup> As might be expected, evidence shows that the effects of disease progression in these patients is linked to a deterioration in HRQoL.<sup>12 13 14 15</sup>

### **2.1.4 Epidemiology**

The company provides an overview of the incidence of kidney cancer in the UK, mostly based on data reported by Cancer Research UK and the National Office of Statistics. Figures of new cases of kidney cancer for England in the CS are cited for 2015, with 9023 new cases (ICD-10 C64 malignant neoplasm of kidney, except renal pelvis), equating to an age-standardised rate of 24.3 per 100,000 in males and 12.3 per 100,000 in females. More recent data identified by the ERG by the Office for National Statistics in England shows that during 2016, 5823 new cases of kidney cancer for males and 3392 for females were recorded (an increase of over 2%), equating to age-standardised rates of 24.5 per 100,000 in males and 12.4 per 100,000 in females.<sup>16</sup> RCC is a sub-type of kidney cancer, accounting for around 80% of all kidney cancer cases, as stated above.

The ERG's quality assessment mostly agrees with that of the company. The ERG disagrees with the company that there is no risk of bias for random sequence generation and for allocation concealment. In the ERG's view the risk is unclear as adequate information has not been provided on procedures. Both the company and the ERG agree that the trial is at a high risk of bias due to being open-label. However, a blinded retrospective review by an independent radiology committee (IRC) was undertaken to minimise detection bias for the PFS and response outcomes in the company's updated analysis. Overall, the ERG is of the opinion that the CABOSUN trial appears to have been well conducted though with some limitations as outlined above.

### **3.1.5 Description and critique of company's outcome selection**

The outcomes in the CS match those listed in the NICE scope and the decision problem.

These are:

- PFS - defined as the interval between randomisation and first documentation of disease progression, or death from any cause. This outcome was originally investigator-assessed. For the regulatory submission, a blinded, retrospective central review of the radiographic images was carried out by an IRC to determine progress and response. The CS presents IRC-assessed results for this outcome. Progression was assessed according to RECIST 1.1 at screening and every two treatment cycles (i.e. every 12 weeks).
- OS - defined as time from randomisation to death from any cause.
- ORR - defined as the proportion of patients at the time of data cut-off with a best overall response of CR (complete response) or PR (partial response), confirmed by a subsequent visit  $\geq 28$  days later (assessment as for PFS).
- Adverse events - graded by Common Terminology Criteria for Adverse events (CTCAE) version 4. Safety was assessed on a schedule based on the date of the first dose, days 15 and 29 of Cycle 1 and 2, and day 1 of each subsequent cycle.

The above outcomes are valid and appropriate endpoints used in cancer trials. Of these, only ORR is not used in the economic model of the CS.

In addition to the listed outcomes, the company states 'Duration of response' under 'all other reported outcomes' (CS Table 8). No definition for this outcome is provided.



HRQoL data were not collected in the CABOSUN trial and hence not reported for the clinical effectiveness section of the CS. Phase II clinical trials generally do not assess outcomes such as HRQoL. HRQoL in cancer trials is an important outcome that should be included, as it generally reflects a patient's day-to-day functioning.<sup>33</sup> For the economic model, the company used other published sources of HRQoL data, as discussed in section 4.3.5 of this report.

### **3.1.6 Description and critique of the company's approach to trial statistics**

The CS reports results for all of the outcomes specified in the NICE scope, apart from HRQoL which had not been assessed in the CABOSUN trial (CS Table 1).

The statistical analysis approaches employed in the CABOSUN trial are summarised in CS Table 11. The CSR states that the statistical analysis plan for CABOSUN is available in an Appendix of the CSR; this was not available to the ERG and was requested by the ERG from the company (clarification question A20).

#### **3.1.6.1 Statistical analysis approaches**

Two different analysis approaches were employed in the CABOSUN trial:

- the original analysis, as reported in the CSR and the trial publication;<sup>24</sup>
- an updated analysis that was conducted by the company to meet regulatory requirements (CS Table 7).

The CS states that the company's submission to NICE is based on the updated analysis and therefore results as reported in the CS differ in some respects to those reported in the trial publication (CS section B.2.2).<sup>24</sup> Results of the updated analysis are also reported in the CSR and in a conference presentation.<sup>30</sup>

Standard statistical methods were used to compare time-to-event outcomes between cabozantinib and sunitinib (CS section B.2.4). Kaplan-Meier (K-M) curves are presented in CS Figure 5 for PFS and in CS Figures 6 and 7 for OS. The hazard ratios were estimated based on Cox regression with a 2-sided log-rank test stratified by IMDC risk group (poor, intermediate) and bone metastases (yes, no) (for a definition of the IMDC risk factors see section 2.1.4). The CS clearly reports the number of patients at risk at each time point; the number of patients censored for in each trial arm, with reasons (CS Table 12 for PFS; CS Table 13 for OS); the median PFS and OS with 95% confidence interval for each trial arm; the hazard ratio (HR) with 95% confidence interval; and the p-value from the log-rank test (CS Figure 5 for PFS; CS Figures 6 and 7 for OS).

The ERG notes that the distribution of patients across risk categories for these two instruments in this trial are broadly similar. Expert clinical advice to the ERG is that MSKCC and IMDC are similar, thus differences between the trials in how patients were classified would be unlikely.

The CABOSUN trial included only patients at intermediate or poor RCC risk, whilst the COMPARZ study included patients with favourable, intermediate and poor risk classifications. The distribution of patients between risk classifications is therefore different between the two trials. Approximately 80% of patients in the CABOSUN trial were at intermediate risk, compared to approximately 54% to 56% in COMPARZ, and approximately 19% of patients were classified as poor risk in CABOSUN compared to 17 to 19% in COMPARZ (all figures based on the IMDC risk classification). The percentage of patients with favourable risk in COMPARZ was 25%, with no favourable risk patients in CABOSUN for the reason stated above. The patient RCC risk profile in COMPARZ is therefore more favourable than in CABOSUN. The CS does not comment on the impact of this difference, but the ERG considers this would likely under-estimate the relative effectiveness of cabozantinib compared to pazopanib in the ITC since patients in the COMPARZ trial overall have a lower RCC risk and accordingly could be expected to respond more favourably to treatment.

Cancer performance status was reported by ECOG classification in CABOSUN and the Karnofsky index in COMPARZ. In CABOSUN around 46% of patients were classified as ECOG 0 (which indicates the patient is fully active, and able to carry on all pre-disease performance without restriction), and around 41% were classified as ECOG 1 (which indicates mild restriction in ability to carry out physical activity and work). In COMPARZ around 75% of patients had a Karnofsky score of 90 to 100%, indicating normal activity, no/minor signs of disease (NB. The data for Karnofsky performance status 70 to 80 and 80 to 100 are the wrong way round in CS Appendix Table 11). An ECOG performance status of 0 is considered comparable to Karnofsky score of 90% to 100%, and an ECOG performance status 1 is comparable to a Karnofsky score 70% to 80%.<sup>37</sup> Thus, the two trials are broadly comparable in terms of cancer performance status, though it appears that a greater proportion of COMPARZ patients were classified as having the highest performance status. Expert clinical advice to the ERG is that this is likely to be due to some of the patients in COMPARZ having favourable risk status (ECOG performance status is one of the constituent variables in the risk status assessment).

Method 3 is presented as an additional analysis to explore comparative treatment effects in RCC risk groups. It does not assume proportional hazards and does not inform the economic model. We provide a brief description and critique of this analysis in Appendix 9.3.

The following sub-sections describe and critique, in turn, methods 1 and 2.

#### **3.1.7.4 ITC: comparison of parametric survival curves**

The CS reports use of a Bayesian statistical method described by Ouwens et al (2010) as a method for conducting an ITC.<sup>36</sup> This method was developed as an alternative to methods of assessing treatment effects which assume proportional hazards. The application of a constant HR implies the assumption that the treatment only has an effect on the scale parameter of a distribution. The method devised by Ouwens et al<sup>36</sup> uses parametric survival distributions to extrapolate outcomes which can be described by two parameters (shape and scale). The time-varying HR is expressed as a difference in scale and a difference in shape of the hazard functions of compared interventions. Ouwens et al<sup>36</sup> consider that encompassing treatment effects on both shape and scale is a more flexible approach to model relative survival. The method can be applied to pairwise meta-analysis of survival curves as well as multiple indirect comparisons of interventions. The similarity and consistency assumptions need to be fulfilled as they would do in other types of indirect comparison (see below).

The method can be used with both individual patient data and aggregated data from Kaplan-Meier curves. Scanned survival curves can be divided into multiple consecutive intervals over the trial follow-up period, and extracted survival proportions can be used to calculate the incident number of progression events or deaths for each interval and patients at risk at the beginning of the interval.<sup>36</sup>

Five parametric models were used by the company in the application of this method, four of which assumed two-parameter distributions (Weibull, log-logistic, log-normal, Gompertz), and one which used a one-parameter (exponential) distribution. The CS states that the exponential model was chosen because it made the same assumption as the previous method of hazard proportionality and allowed comparison. Model fit was assessed using the deviance information criteria (DIC) (CS Table 23).

Bayesian models were fitted using sunitinib as the reference treatment, and estimated treatments in terms of their effect on the reference parameters. The CS states that effect transitivity is an underlying model assumption. The transitivity assumption (also known as

- **Plausibility of extrapolation:** The company states that visual inspection of the curves by clinical oncologists led to the conclusion that the lognormal, loglogistic and gamma distributions give unrealistically optimistic long-term survival.

We show the fitted curves together with CABOSUN KM data in Figure 15 and selected summary statistics in Table 13 below. The ERG agrees that the exponential has a reasonable visual and statistical fit for both treatments and that it yields plausible estimates of long-term survival: 13% at five years for sunitinib in comparison with 21% for an observational cohort from the IMDC dataset that includes patients with a better risk profile.<sup>47</sup> Use of an exponential distribution for both treatments conflicts with the conclusion that OS hazards are not proportional. But we suggest that the exact shape of the CABOSUN KM curves should not be over-interpreted given the modest sample size (n=157) and lack of explanation for why the curves should come together and then diverge between about 13 and 20 months. The Weibull distribution and Gompertz provide reasonable alternatives for scenario analysis.

The ERG is concerned that the OS curves appear to have been fitted to CABOSUN January 2017 data cut, rather than the most recent July 2017 dataset which was less favourable for cabozantinib (CS B.2.6 Figures 6 and 7). The CS stated that the data cut off January 2017 was used (CS page 39, Table 13 and Figure 6). The KM plot was reproduced in the economic chapter (CS B.3.3 Figure 13) and the KM data provided by the company in response to a clarification question also related to this cut-off. Failure to use the most recent available data will introduce bias in favour of cabozantinib. We consider this issue in ERG additional analysis; section 4.4.1 below.

### **OS indirect comparisons**

Figure 16 shows the ITC parametric and best-fitting FP survival curves in relation to the CABOSUN KM curves. We omit the COMPARZ KM curves from these graphs for clarity; but note that they are similar to the CABOSUN KM curve for cabozantinib and lie above the CABOSUN KM curve for sunitinib. This reflects the better risk status of participants in COMPARZ than in CABOSUN. The summary OS statistics are in Table 13 below.

The company use a second order FP model with  $P1=P2=-1$  for OS in their ITC base case and three random effect parametric curves (exponential, Weibull and Gompertz) and two FPs ( $P1=-0.5$ ,  $P2=0$ ) and ( $P1=-1$ ,  $P2=0$ ) in scenario analysis. Their rationale for this choice is outlined in the CS:

**Table 26 ERG preferred assumptions and scenarios**

	<b>Preferred assumptions</b>	<b>Scenarios</b>	<b>Reason for analysis</b>
Time horizon	20 years	5/ 10 years	Reflects full lifetime, but with scenario analysis to show impact of extrapolation
Persistence of OS and PFS benefit	5 years from baseline	10/ 20 years	Given the weakness of evidence for the OS difference, we take a conservative approach, with progression and mortality hazards for cabozantinib equal to those of sunitinib after 5 years (3 years after trial follow up).
OS curves	Simple indirect comparison	HR = 0.74 (Jan 2017 analysis). And no effect (HR=1)	Exponential OS for sunitinib (separate fit to CABOSUN). Cabozantinib estimated from sunitinib curve and HR=0.80 (July 2017 CABOSUN update). OS assumed equal for pazopanib and sunitinib, based on the relationship in COMPARZ. Exploratory scenarios to compare with company model and assess impact of OS.
	Age-related mortality		Minimum mortality rate based on general population life table (ONS 2014-16).
PFS curves	Lognormal direct comparison	Exponential and Gompertz	Same as in company direct base case. Lognormal gives most plausible fit, and we use selected alternatives for scenarios (see table below).
TTD curves	Lognormal direct comparison	All available	We agree that the lognormal gives the best fit, but there is little reason to choose between other functions, so we use all in scenario analyses.
Health state utilities	PF and PD utilities from Tivozanib TA512 (base case)	Swinburn, Pazo TA215 and Suni TA169	We follow the company approach, with the utilities for pre and post-progression based on values accepted by committee for tivozanib, with scenarios testing alternative sources.
AE disutilities	Amdahl disutility, applied for 4 weeks to TEAE with $\geq 5\%$ incidence	Range of disutilities, 8 week duration and $\geq 2\%$	Again, we follow the company approach, but conduct additional analyses to test the sensitivity of the model to adverse events.

	<b>Preferred assumptions</b>	<b>Scenarios</b>	<b>Reason for analysis</b>
Dose intensities	Dose intensities from CABOSUN (94.3% cabo, 83.9% suni) and 86% for pazo from tivozanib STA	Tested 86% for all first-line drugs, and also 100%	Company's assumptions are reasonable but we explore the impact on costs of uncertainty over dose intensity, using the range suggested by committee considerations from the NICE tivozanib appraisal guidance
Subsequent treatment costs	Use of second-line treatments from trials	Company and ERG scenarios	Utilisation from trials reflects effectiveness evidence, but it includes drugs not recommended or available in UK. The company includes a scenario based on clinical advice, using only NICE recommended second-line drugs. We test 2 other scenarios. ERG 1: equal distribution of NICE approved second-line drugs (20% each drug and 10% BSC; cabozantinib 1st line patients only eligible for nivolumab, everolimus or lenvatinib with everolimus, 30% each drug and 10% BSC). ERG 2: based on clinical advice we assume use only of nivolumab, cabozantinib, lenvatinib with everolimus (30% each drug, and 10% BSC; cabozantinib 1st line patients only eligible for nivolumab and lenvatinib with everolimus, 45% each drug and 10% BSC).
Health state management costs	Based on resource use assumptions from tivozanib appraisal	Company scenario based on clinical advice. More expensive blood test (£20)	Clinical advisors to the ERG agreed that resource use assumptions were appropriate
Adverse event costs	Series of assumptions based on clinical advice and guidance.		As above
Age of cohort	62.8 years	55/75 years	Exploratory: to assess applicability to the UK RCC population

Table 1 ERG approach to modelling treatment effects

	Company base case (scenarios)	Comments	ERG preferred assumptions
OS curves	<p><b>Direct:</b> Exponential (Weibull &amp; Gompertz)</p> <p><b>ITC:</b> FP model with <math>P1=P2=-1</math> (exponential; Weibull; Gompertz; and FP <math>P1=-0.5</math>, <math>P2=0</math> &amp; <math>P1=-1</math>, <math>P2=0</math>)</p>	<p>CABOSUN is not powered for OS and data are relatively immature, so the KM curves are noisy. Reason for crossover is unclear. Uncertainties over the ITCs due to differences in trial populations.</p> <p>Given these reservations, the exponential, Weibull and Gompertz are reasonable for the direct analysis. For the ITC, the exponential and FP <math>P1=P2=-1</math> curves are reasonable. But other scenarios predict unrealistic long-term survival. Fitted curves based on Jan 2017 CABOSUN data, rather than less favourable July 2017 dataset.</p>	<p>Simple indirect comparison assuming:</p> <ul style="list-style-type: none"> <li>• Sunitinib OS curve based on company's exponential fit to CABOSUN;</li> <li>• Cabozantinib calculated from sunitinib curve and HR from July 2017 CABOSUN results;</li> <li>• Pazopanib curve assumed equal to sunitinib (based on COMPARZ results).</li> </ul>
PFS curves	<p><b>Direct:</b> lognormal (Exponential, Weibull &amp; Gompertz)</p> <p><b>ITC:</b> FP <math>P1=P2=-1</math> (exponential, Weibull and Gompertz)</p>	<p>CABOSUN PFS analysis is more mature. ITC is subject to uncertainty due to differences in trial populations, unclear if similarity assumption is met.</p> <p>Direct comparisons with lognormal, exponential and Gompertz are reasonable, but the Weibull has poor visual fit. For ITC, Lognormal and loglogistic models give best balance of fit and extrapolation.</p>	<p>Simple indirect comparison: use lognormal separately fitted to CABOSUN for cabozantinib and sunitinib and assume equivalence for pazopanib and sunitinib (COMPARZ). We also test alternative separately fitted curves: exponential and Gompertz curves.</p>
TTD curves	<p><b>Direct:</b> lognormal (exponential, Weibull, Gompertz &amp; gamma).</p>	<p>TTD data are mature, with little difference in the visual fit or extrapolation of survival functions. There is no obvious reason for excluding the loglogistic from scenario analysis. The assumption of equal TTD for pazopanib and sunitinib is reasonable given similarity in COMPARZ.</p>	<p>Lognormal for base case, and all other distributions in scenario analysis.</p>

## 5 End of life

The CS argues that cabozantinib meets the NICE end-of-life criteria. Table 35 (CS Table 28) summarises their justification for reaching this conclusion.

**Table 35 End-of-life criteria (CS Table 28)**

<b>Criterion</b>	<b>Data available</b>
<b>The treatment is indicated for patients with a short life expectancy, normally less than 24 months</b>	In the IMDC validation study (1028 patients receiving first line VEGF-targeted treatment for metastatic RCC), median OS from the start of treatment was 22.5 months (18.7-25.1) in the intermediate risk group and 7.8 months (6.5-9.7) in the poor risk group.
<b>There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment</b>	In the CABOSUN trial, median survival was 30.3 months (95% CI 14.6, NE) in the cabozantinib arm vs. 21.0 months (95% CI 16.3, 27.0) in the sunitinib arm, an estimated 9.3 - month difference in the medians at a median follow-up of 28.9 months. In the economic modelling, which extrapolates beyond the duration of the trial, cabozantinib is associated with a gain of 0.66 life years (7.9 months) compared with sunitinib. The other treatment currently used in the NHS is pazopanib. Pazopanib was found to have similar efficacy to sunitinib in terms of both PFS and OS in a head-to-head trial in 1110 patients with previously untreated metastatic RCC (Motzer 2013). In the economic modelling, cabozantinib is associated with a gain of 0.80 life years (9.6 months) compared with pazopanib.

The ERG's analysis confirms that cabozantinib offers an additional extension of life, which exceeds 3 months when compared to sunitinib or pazopanib (5.9 months in ERG's analysis). However, the submitted CS model and results from the ERG's preferred assumptions give mean OS estimates exceeding 24 months for sunitinib and pazopanib (■ life years without discounting in the ERG analysis). We are therefore of the opinion that cabozantinib does not fully meet the NICE criteria for being considered as a life-extending treatment for people with a short life expectancy.

## 5 Innovation

The CS suggests that the superior effectiveness compared with current treatments can be explained by its novel mechanism of action. Cabozantinib is the first and only multi-targeted therapy for RCC which targets pathways involved in both tumour growth and drug resistance