NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Cabozantinib for untreated advanced renal cell carcinoma [ID1208]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Cabozantinib for untreated advanced renal cell carcinoma [ID1208]

Contents:

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

Redacted

Pre-meeting briefing

Cabozantinib for untreated metastatic renal cell carcinoma

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting

1

Disease background and management

Kidney cancer

- More common in men than women
- Five-year survival is 56%, varying with age
- 86% of renal cancers are renal cell carcinoma

Renal cell carcinoma

- Estimated 9,045 new diagnoses in England per year
- Disease is often locally advanced or metastatic at point of diagnosis
- Early stage disease can be treated surgically – half of patients who have surgical treatment will develop metastatic disease
- Overall survival for people with metastatic disease is 8 months to 3.6 years

2

IMDC risk categories of prognosis

IMDC 2013	
Factor	**Poor prognostic factor**
Kafnofsky Performance Status	Less than 80%
Time from diagnosis to treatment	Less than 12 months
Anaemia	Lowered haemoglobin count
Hypercalcemia	Elevated calcium concentration
Neutrophilia	Elevated neutrophil count
Thrombocytosis	Elevated platelet count

Risk categories

Favourable risk
No factors

Included in this appraisal

Intermediate risk
One or two factors

Poor risk
Three or more factors

IMDC, International Metastatic RCC Database Consortium

3

Source: Heng et al, *Lancet Oncol*, 2013;14(2):141–148
Company submission, p14–15, section 1.3

Cabozantinib (Cabometyx)

Anticipated UK marketing authorisation	The treatment of advanced renal cell carcinoma (RCC) in treatment-naive adults with intermediate or poor risk per IMDC criteria
Administration	Oral
Mechanism of action	Multiple receptor tyrosine kinase inhibitor, including hepatocyle growth factor receptor protein (MET), vascular endothelial growth factor (VEGF) and AXL. Targets pathways implicated in tumour progression, angiogenesis, pathologic bone remodelling, and drug resistance.
Dosage	60 milligrams (one tablet) cabozantinib once daily for 40 and 20 milligram tablets are available so that the dose can be reduced if necessary
PAS	Yes, a simple PAS has been agreed with the DHSC

4

Source: company submission p11-12, section 1.2, table 2.

Decision problem

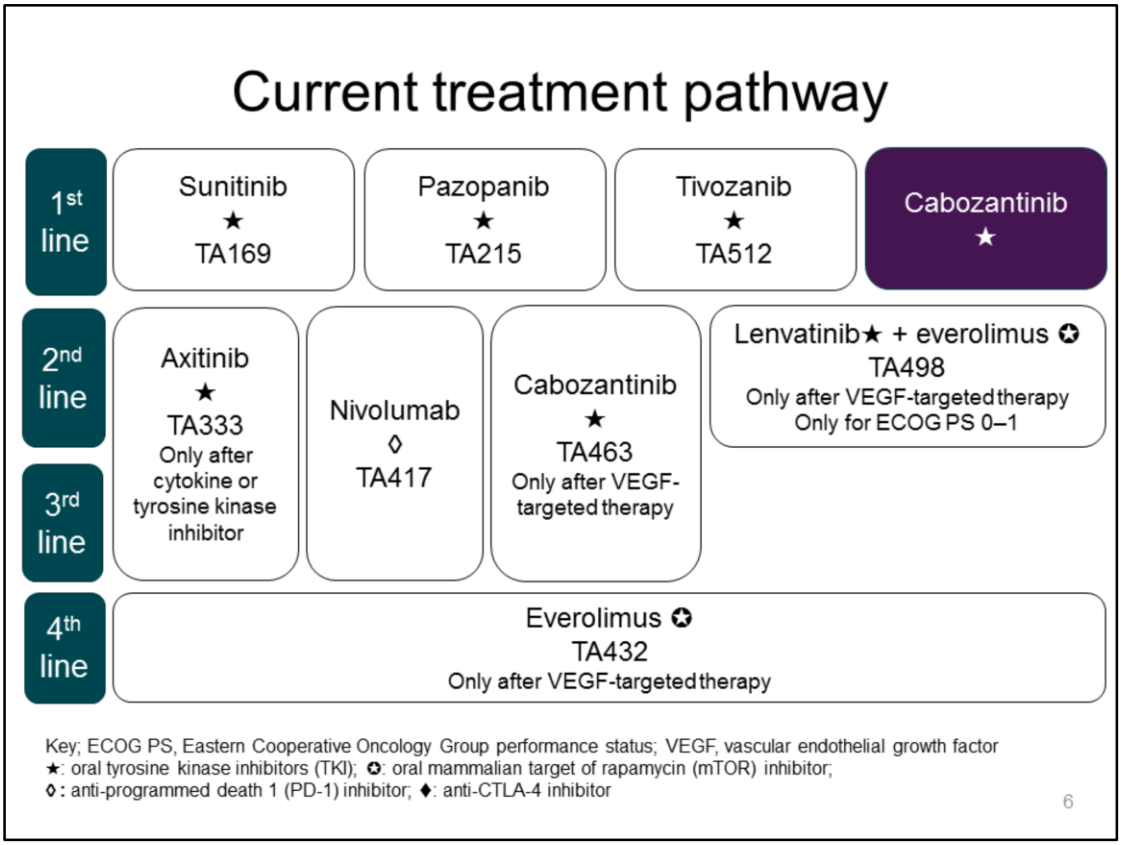
	Final scope issued by NICE	Company's decision problem
Population	People with untreated, intermediate or poor risk (as per IMDC), locally advanced or metastatic renal cell carcinoma	As per the scope
Comparators*	Pazopanib Sunitinib	As per the scope
Outcome	• Overall survival • Progression-free survival • Response rates • Adverse effects of treatment • Health-related quality of life	• Overall survival • Progression-free survival • Response rates • Adverse effects of treatment

Company states that quality of life data were not collected in the phase II (CABOSUN) trial and data from published sources were used in the model

* Tivozanib not recommended at time of scoping

5

Source: company submission p10, table 1



VEGF-targeted therapy includes tyrosine kinase inhibitors

Comments from patient groups

- Patient groups
  - People may experience constant pain as well as psychological effects e.g. depression, loss of confidence and self-worth
  - Many patients have to give up work due to debilitating effects of disease and treatments available – leads to financial pressures
  - Few treatment options available currently and adverse effects are significant e.g. extreme fatigue, severe hand and foot syndrome, chronic diarrhoea
  - No biomarkers to predict who will respond to each drug, therefore range of treatment options important
  - Cabozantinib could be used to address an area of significant unmet need for an effective first-line treatment for bone metastases

7

Joint submission from NCRI, ACP and RCP
Submission from KCSN

Comments from professional groups

- Professional groups
  - Pazopanib or sunitinib currently used first-line
  - Clinicians are experienced with using cabozantinib in second-line
  - Adverse event profile of cabozantinib is similar to other tyrosine kinase inhibitors
  - Some uncertainty in magnitude of benefit of cabozantinib over sunitinib due to small size of the CABOSUN trial
  - The performance of both trial arms were a little poorer than expected

8

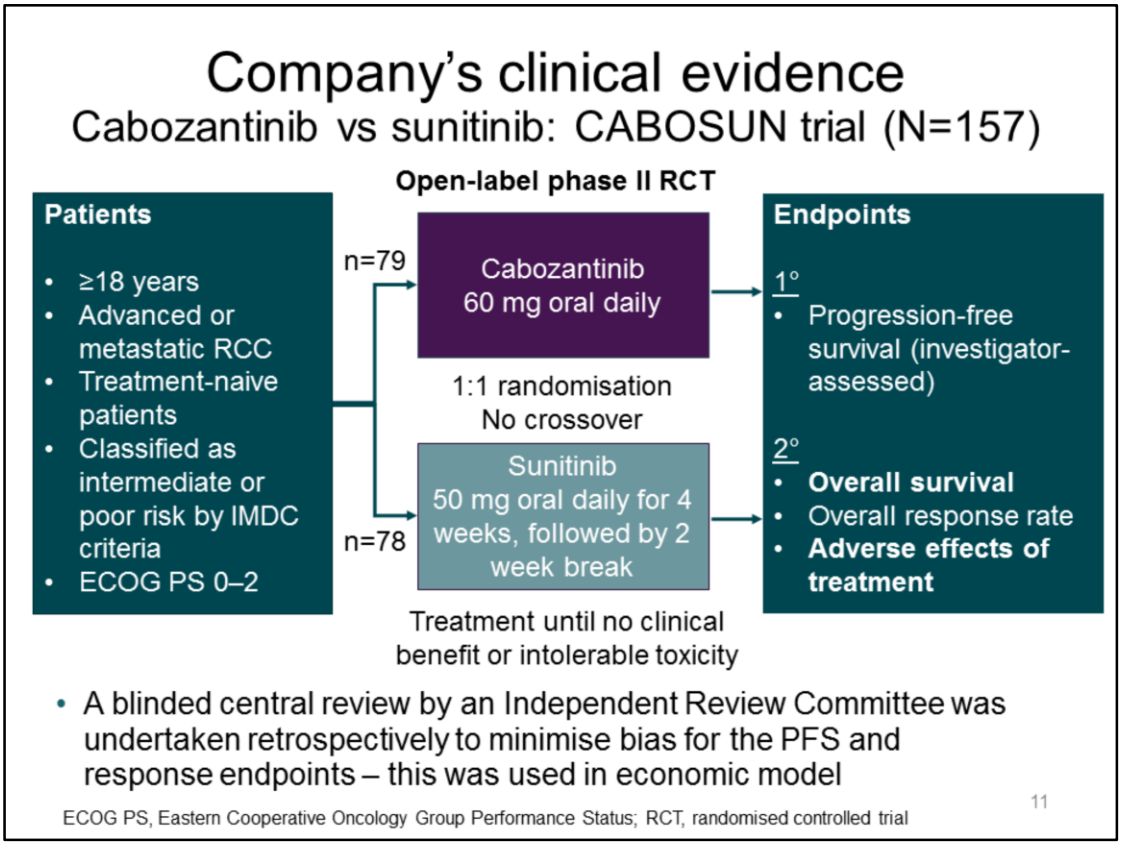
Joint submission from NCRI, ACP and RCP
Submission from KCSN

Clinical effectiveness evidence

Company submission section 3.2

Key clinical evidence

- CABOSUN Phase II randomised controlled trial (RCT) compares cabozantinib with sunitinib
  - Retrospective analyses conducted using FDA approved censoring rules and Independent Review Committee analyses of PFS (slide 12)
- Company conducted an indirect treatment comparison to compare cabozantinib with pazopanib (slide 22)
  - Utilised data from CABOSUN and the COMPARZ trial, an RCT that compares sunitinib with pazopanib
  - Results from indirect treatment comparison may be biased by difference in populations between the 2 trials (slide 23)
- Company and ERG agree that CABOSUN and COMPARZ trial were well designed and conducted, with a low risk of bias for most domains



Source: company submission p29–32, Section 2.3
Primary manuscript: Choueiri et al, *J Clin Oncol* 2017;35(6):591–597

Notes
- No UK patients (all patients US)
- Randomisation stratified by by IMDC risk group (intermediate or poor) and presence of bone metastasis (yes/no)

CABOSUN data cuts and censoring

	Original report (Choueiri 2016)	Clinical Study Report and Choueiri 2017	
Cut-off date (PFS and ORR)	April 2016	**September 2016***	
Cut-off date (OS)	April 2016	**January 2017 (CSR)*** July 2017 (Choueiri 2017)	
Censoring rules (PFS)	Alliance	FDA guidance	
Censor for non-protocol systemic anticancer therapy	No	Yes	
Censor if event after ≥2 missing assessments	No	Yes	
Stratified analysis†	Yes	Yes	
p-value sided	1	2	
Reader	Investigator	Investigator	Independent
No. of patients with radiographs	157	157	156
No. of events	123	107	92

*Data cuts in bold were used in the economic model
†Stratification factors were IMDC risk group (poor, intermediate) and bone metastases (yes, no)

Source: Company submission p28, section 2.2. table 7

Notes

Some discrepancies between original report and data used in model as FDA censoring rules were applied in the CSR. Data from the CSR was used in regulatory submissions and in the economic model

Analyses by blinded IRC-assessed PFS and ORR was conducted post-hoc as basis for the regulatory submission to reduce risk of bias. IRC-assessed PFS was used in the economic model.

one site with one subject in the sunitinib arm declined to participate in radiographic image collection).

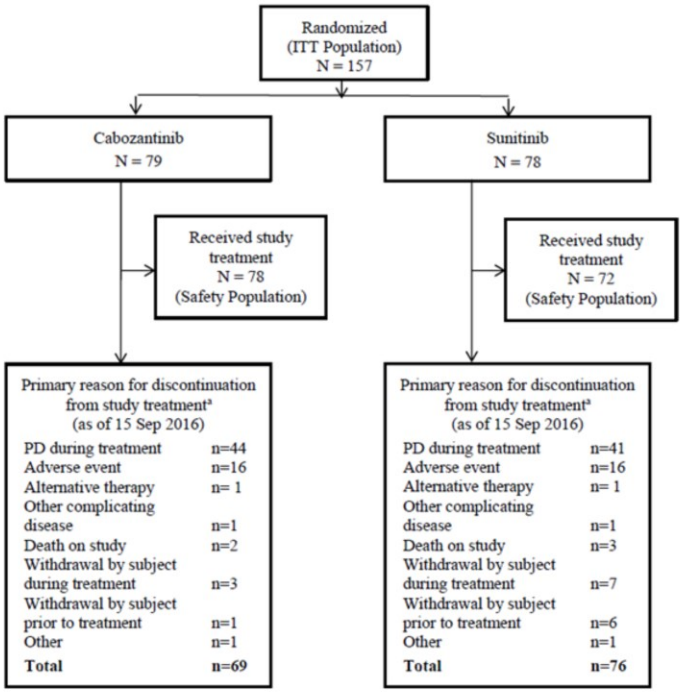
Statistical analysis

- Original sample size and power calculation:
  - With 123 events (progressions or deaths)
  - 85% power to detect an HR of 0.67 for PFS assuming a one-sided type I error of 0.12
  - Equivalent to increase in median PFS from 8 months for sunitinib to 12 months for cabozantinib
- Analysis presented in submission uses FDA-recommended censoring rules and two-sided p-values
  - Sensitivity analyses were performed to explore effect of potentially informative censoring – showed PFS results robust
  - ERG: using this approach the primary outcome of PFS is statistically underpowered
    - 2-sided test has less statistical power
    - 92 events occurred which is less than the planned target of 123 events
  - Company: OS data were immature at this data cut-off and there was a notable degree of censoring around the median estimates

13

Source: Company submission. p34-35, section 2.4, table 11
ERG report p42-44, section 3.1.6

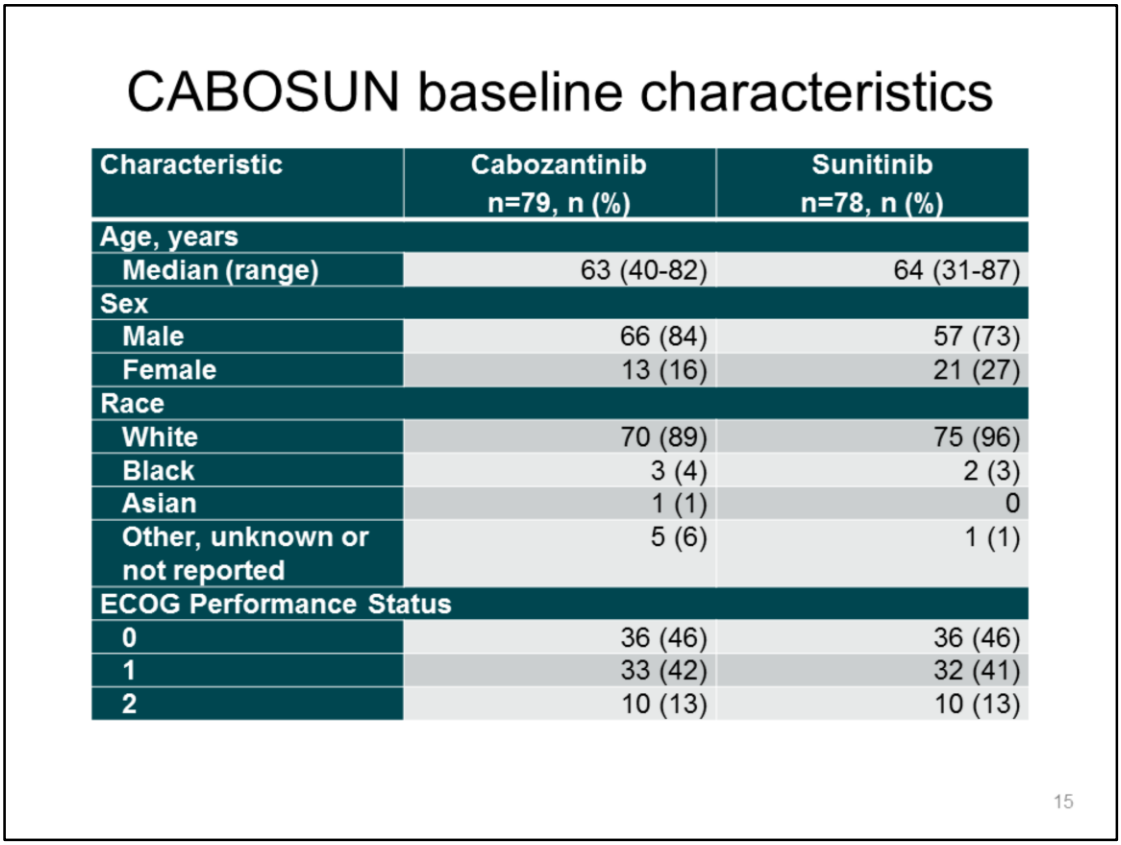
Participant flow in CABOSUN



- ERG: unexplained imbalances between trial arms in missing data on tumour assessments and in patient discontinuations due to withdrawal of consent
- Unclear whether these might have introduced bias

14

Source: Company submission appendices, p217, D1.2, figure 52
ERG report p45-46, section 3.1.6.7



CABOSUN baseline characteristics

Characteristic	Cabozantinib n=79, n (%)	Sunitinib n=78, n (%)
Age, years		
Median (range)	63 (40-82)	64 (31-87)
Sex		
Male	66 (84)	57 (73)
Female	13 (16)	21 (27)
Race		
White	70 (89)	75 (96)
Black	3 (4)	2 (3)
Asian	1 (1)	0
Other, unknown or not reported	5 (6)	1 (1)
ECOG Performance Status		
0	36 (46)	36 (46)
1	33 (42)	32 (41)
2	10 (13)	10 (13)

15

Source: Company submission, p33, section 2.3, table 10

CABOSUN baseline characteristics

Characteristic	Cabozantinib n=79, n (%)	Sunitinib n=78, n (%)
IMDC risk group		
Intermediate	64 (81)	63 (81)
Poor	15 (19)	15 (19)
Bone metastases		
Yes	29 (37)	28 (36)
No	50 (63)	50 (64)
Prior nephrectomy		
Yes	57 (72)	60 (77)
No	22 (28)	18 (23)
Metastases		
≥ 1 metastatic site	79 (100)	78 (100)
Visceral metastases	61 (77)	56 (72)

- Clinical expert advice to the ERG suggested that the baseline characteristics were generally representative of patients in the UK
  - One expert noted that a higher proportion of patients with prior nephrectomy
- Clinical experts suggest any differences between arms not large enough to be of clinical importance

16

Source: Company submission, p33, section 2.3, table 1

CABOSUN subsequent therapies

Category	Cabozantinib (N = 79), n (%)	Sunitinib (N = 78), n (%)	Available in NHS
VEGFR-TKI therapies	36 (46)	35 (45)	
Axitinib	18 (23)	15 (19)	Yes
Pazopanib	13 (16)	9 (12)	No
Sunitinib	10 (13)	10 (13)	No
Cabozantinib	1 (1)	5 (6)	Yes
Sorafenib	1 (1)	2 (3)	No
Anti-PD-1/PD-L1 targeting agents	10 (13)	12 (15)	
PD-1 inhibitor	6 (8)	6 (8)	No
Nivolumab	4 (5)	6 (8)	Yes
Other selected systemic therapies			
Temsirolimus	7 (9)	3 (4)	No
Everolimus	6 (8)	15 (19)	Yes
Bevacizumab	0	5 (6)	No
Interleukins	2 (3)	1 (1)	No
Interferons	1 (1)	0	No

- Data used to inform economic model
- Median time to first systemic non-radiation anti-cancer therapy was longer in the cabozantinib arm

17

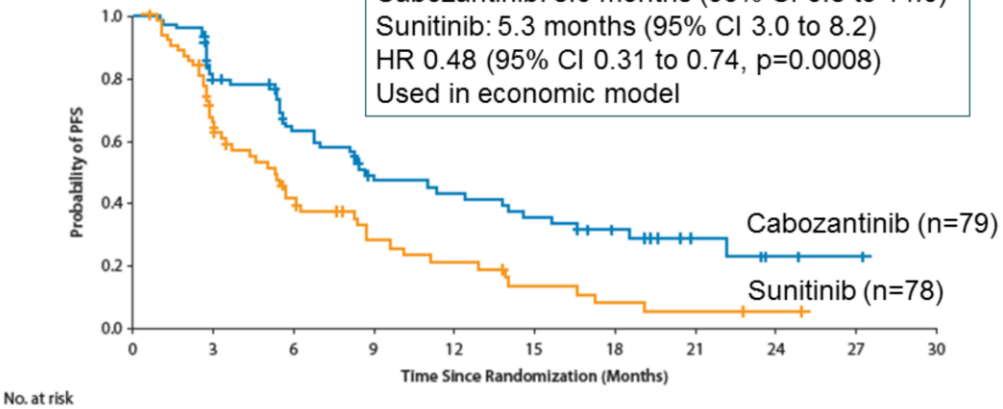
Source: Company submission, table 50 appendix L

Progression-free survival results

Post-hoc IRC assessed (Sep 2016 data cut)

PFS by IRC

Median PFS – FDA censoring rules
Cabozantinib: 8.6 months (95% CI 6.8 to 14.0)
Sunitinib: 5.3 months (95% CI 3.0 to 8.2)
HR 0.48 (95% CI 0.31 to 0.74, p=0.0008)
Used in economic model



No. at risk

Cabozantinib	79	51	37	24	22	18	12	5	2	1	0
Sunitinib	78	36	21	12	9	5	3	2	1	0	0

Data cut-off: September 15, 2016

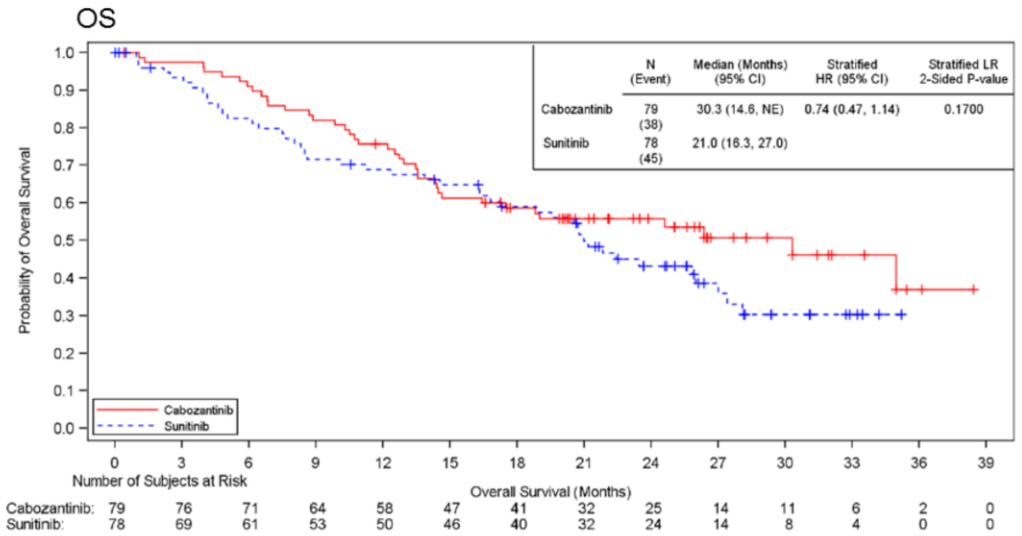
- Results similar to investigator-assessed analysis; 8.3 vs 5.4 months, HR=0.56 (95% CI 0.37 to 0.83, p=0.0042

18

Source: Company submission, p36–38, section 2.6, figure 5

Overall survival results

Jan 2017 data cut



- Company uses Jan 2017 data cut in economic model
- ERG: Impact of subsequent treatment on OS uncertain
- ERG: No explanation of why curves cross in company submission

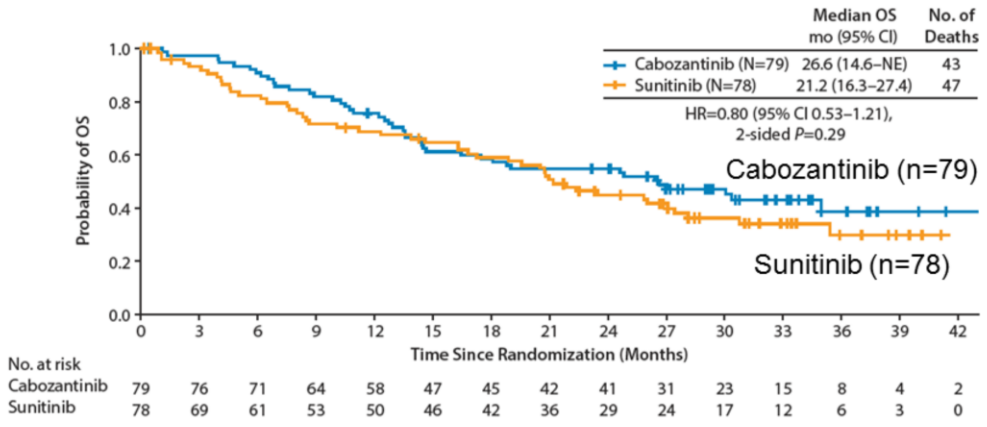
19

Source: Company submission, p39–41, section 2.6, figure

Overall survival results

Jul 2017 data cut

OS



- ERG notes that it is not stated in the company submission why data from the earlier (January 2017) data cut were used in preference
- ERG uses data from the July 2017 data cut in its base case

20

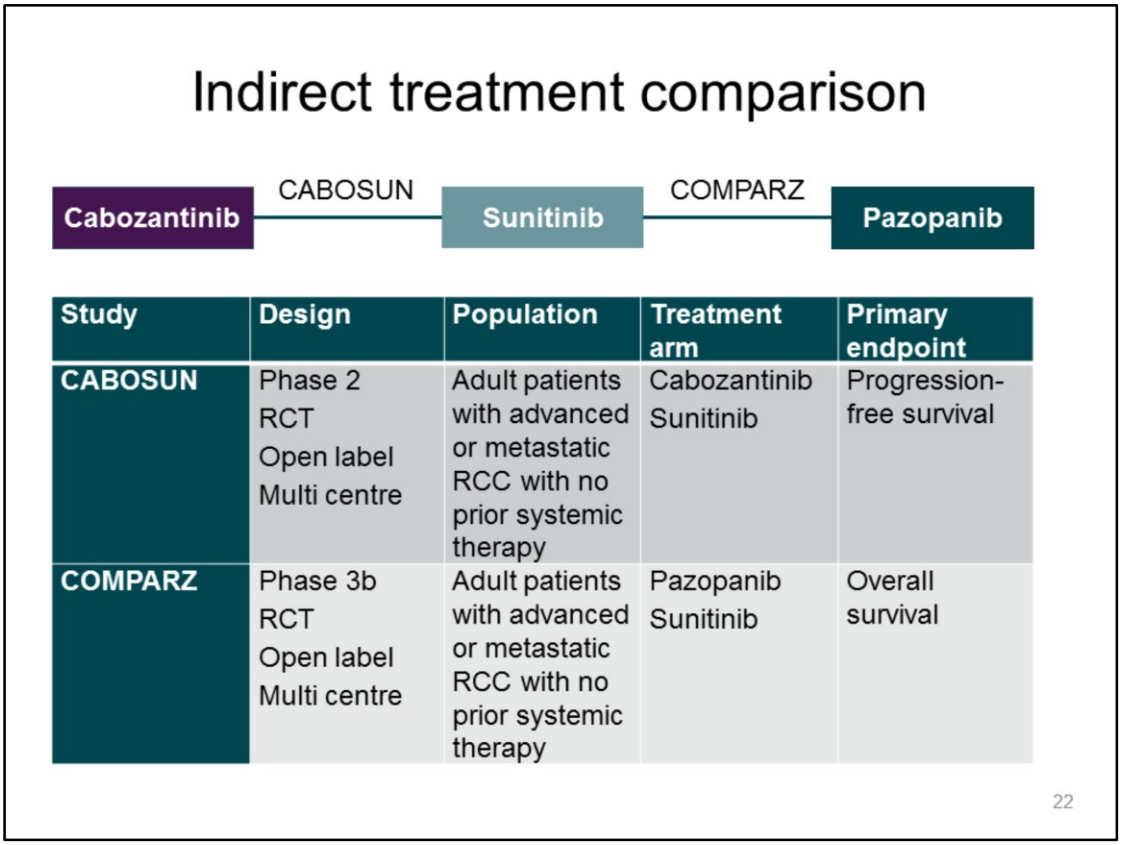
Source: Company submission, p39–41, section 2.6, figure

Subgroup analysis

Survival by risk group

	Cabozantinib Median, months (95% CI)	Sunitinib Median, months (95% CI)	HR (95% CI)
Intermediate/poor	n=79	n=78	
PFS	8.6 (6.8 to 14.0)	5.3 (3.0 to 8.2)	0.48 (0.31 to 0.74)
OS (July 2017)	26.6 (14.6 to NE)	21.2 (16.3 to 27.4)	0.80 (0.53 to 1.21)
Intermediate	n=64	n=63	
PFS	11.4	6.8	0.52 (0.32 to 0.82)
OS	30.3 (16.4 to NE)	23.5 (18.9 to 28.1)	0.80 (0.45 to 1.31)
Poor	n=15	n=15	
PFS	6.1	2.7	0.31 (0.11, 0.92)
OS	18.4 (6.1 to NE)	6.4 (2.2 to 22.4)	0.51 (0.20 to 1.32)

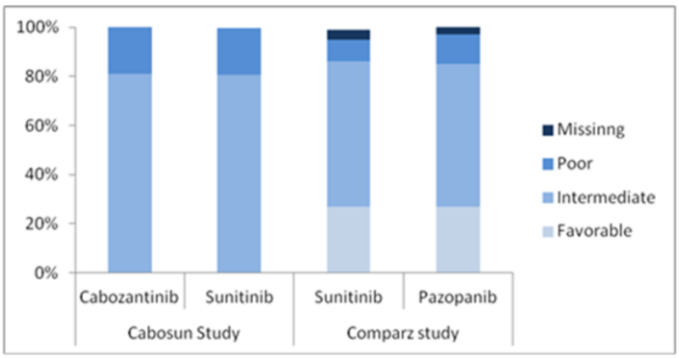
21



Source: Company submission p50, table 19 and p54 figure 11

Indirect Treatment Comparison – Heterogeneity

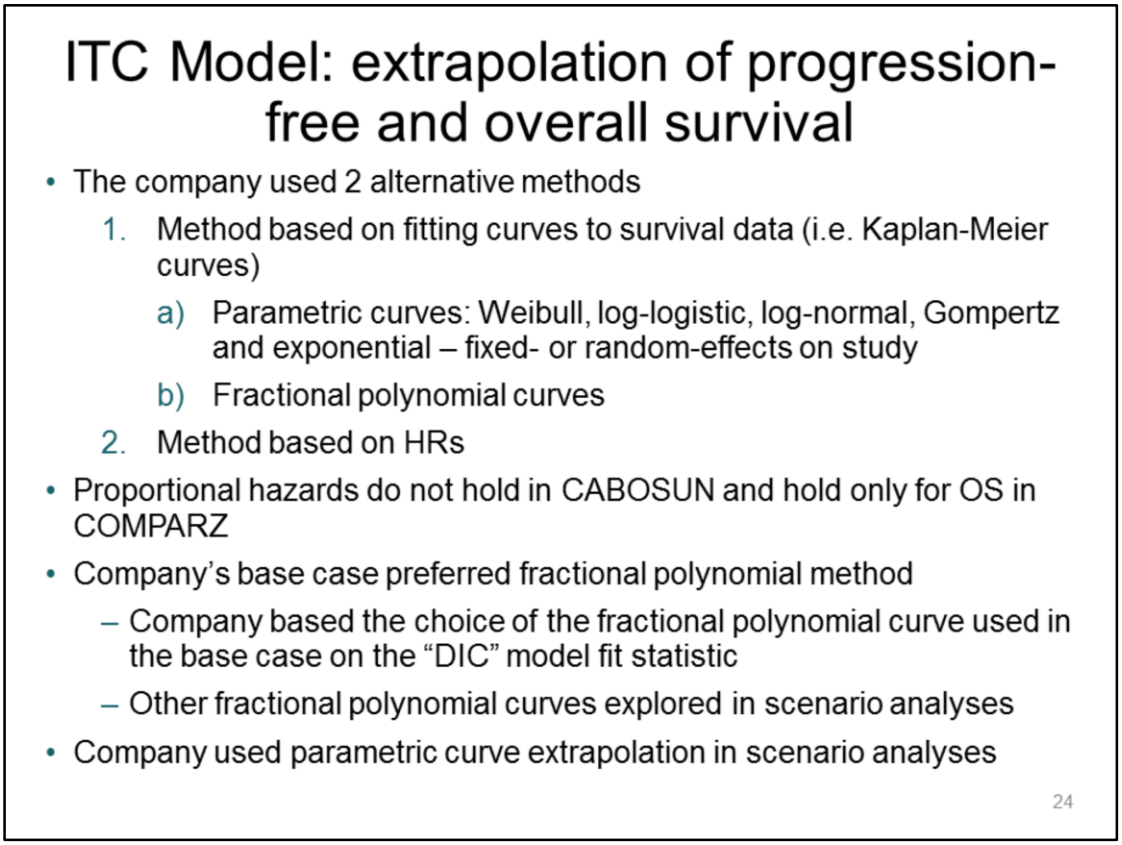
Distribution of risk groups in CABOSUN and COMPARZ



- CABOSUN and COMPARZ have different distribution of risk groups – COMPARZ has 25% favourable risk compared with none in CABOSUN
- More patients had bone metastases at baseline in CABOSUN (36%) compared with COMPARZ (18%)
- ERG: effect of these differences on the ITC results uncertain

23

Source: company submission. P54 figure 10
ERG report



Source company submission: p56, B1.2, table 24
Company submission p88–93, B3.3

To note:
- First order polynomials not considered in both OS and PFS as poorer statistical fit
- The second order curves not considered for OS were considered to have an unreasonably flat tail
- The second order curves not considered for PFS were not considered due to a high PFS rate at year 5

Proportional hazards

- Progression free survival
  - Company states that proportional hazards assumption does not hold for either CABOSUN or COMPARZ
  - ERG: For CABOSUN this conclusion was not supported by the proportionality test or by the Schoenfeld and log-cumulative hazard plots
- Overall survival
  - Company states that proportional hazards assumption does not hold for CABOSUN but holds for COMPARZ
  - ERG: Agrees with company but notes that the exact shape of the CABOSUN Kaplan–Meier curves should not be over-interpreted due to modest sample size and lack of explanation for why the curves come together and then diverge

25

ERG report

ITC Model: fractional polynomial curves

OS				PFS			
1st order fractional polynomial		2nd order fractional polynomial		1st order fractional polynomial		2nd order fractional polynomial	
Model	DIC	Model	DIC	Model	DIC	Model	DIC
P=-1	1722.8	P1=-0.5, P2=0	1716.5	P=-1	1910.4	P1=-0.5, P2=0	1852.1
P=-0.5	1739.5	P1=-1, P2=0	1713.9	P=-0.5	1932.0	P1=-1, P2=0	1840.3
P=0	1757.7	P1=-1, P2=-1	1711.9	P=0	1945.9	P1=-1, P2=-1	1825.0
P=0.5	1769.0	P1=-1, P2=0.5	1716.2	P=0.5	1947.6	P1=-1, P2=0.5	1850.4
P=1	1773.0	P1=-1, P2=1	1718.3	P=1	1943.6	P1=-1, P2=1	1858.1

Key: = used in base case = used in scenario analysis

- Company and ERG also considered parametric curves in scenario analysis – further details in cost effectiveness section

26

Source company submission: p56, B1.2, table 24
Company submission p88–93, B3.3

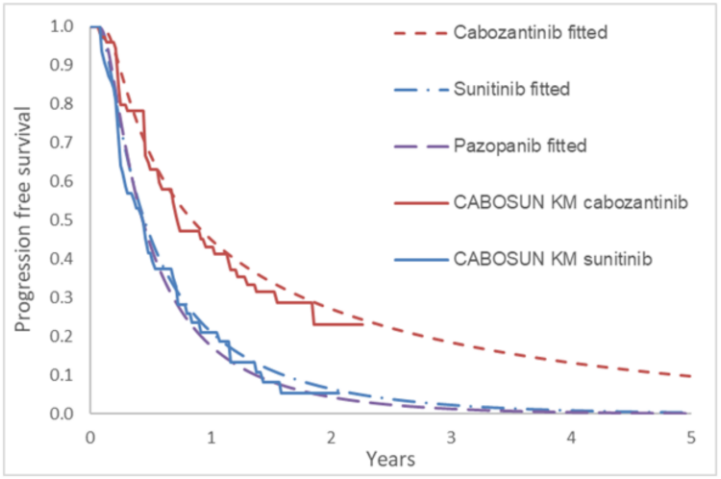
To note:

- First order polynomials not considered in both OS and PFS as poorer statistical fit
- The second order curves not considered for OS were considered to have an unreasonably flat tail
- The second order curves not considered for PFS were not considered due to a high PFS rate at year 5

ITC Model: extrapolation of PFS

Company prefers fractional polynomial distribution

Fitted PFS curve with 2nd order fractional polynomial (P1=P2=-1)



- ERG: Good fit to data but long term projections optimistic
  - Lognormal and log logistic provide best balance of visual/statistical fit and long term projection of PFS

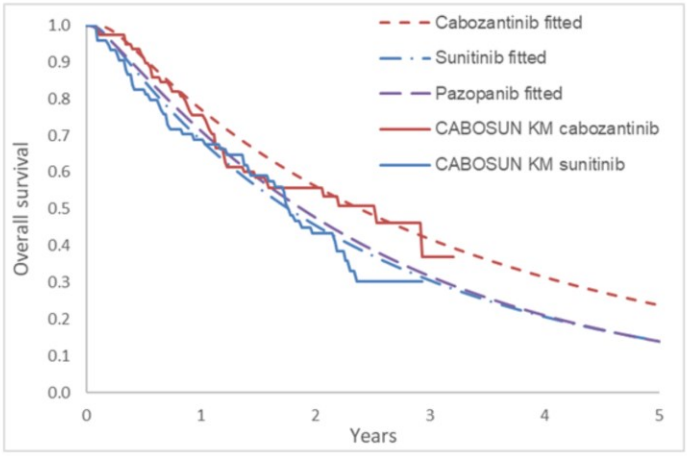
27

Source: ERG report, p90, figure 18F

ITC Model: extrapolation of OS

Company prefers fractional polynomial distribution

Fitted OS curve with 2nd order fractional polynomial (P1=P2=-1)

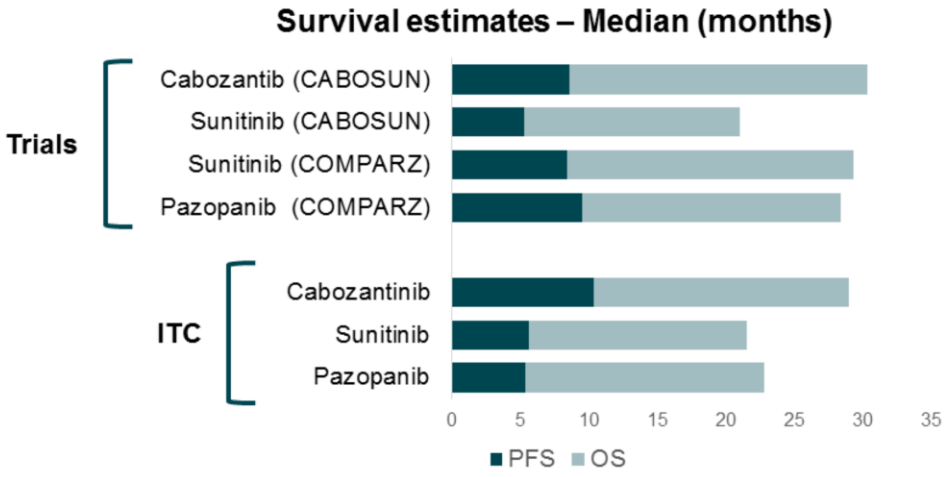


- ERG: Small sample size, immature OS data and heterogeneity in ITC → difficult to assess fit and extrapolation
- ERG: provides reasonable fit and provides plausible long-term survival estimates
  - Random effects exponential also reasonable

28

Source: ERG report p85, Figure 16F

ITC results for company's preferred method (fractional polynomial curves)



- Time varying hazard ratio from COMPARZ data was applied to CABOSUN efficacy data to generate pazopanib ITC data -> median PFS and OS in ITC shorter than COMPARZ but relative relationship maintained over time

29

Adverse events in CARBOSUN

Summary

	Cabozantinib n=78, n (%)	Sunitinib n=72, n (%)
AE	75 (96)	71 (99)
Related AE	74 (95)	70 (97)
Worst AE, grade 3 or 4	53 (68)	47 (65)
Worst related AE, grade 3 or 4	47 (60)	45 (63)
Grade 5 AE up to 30 days after last dose of study treatment	3 (3.8)	6 (8.3)
Grade 5 AE > 30 days after last dose of study treatment	1 (1.3)	3 (4.2)
Related grade 5 AE at any time	2 (2.6)	4 (5.6)
Serious AE	38 (49)	37 (51)
Related serious AE	28 (36)	26 (36)
Deaths	38 (49)	43 (60)
Death up to 30 days after last dose of study treatment	4 (5.1)	8 (11)
Death > 30 days after last dose of study treatment	34 (44)	35 (49)
Discontinuation of study due to AE	21%	22%

30

Source: Company submission, p55, section 2.10, table 25

ERG comments on clinical effectiveness

Performance of sunitinib arm

- ERG noted that the PFS of 5.3 months for sunitinib in CABOSUN was lower than achieved in other clinical trials

	CABOSUN	CheckMate 214	Motzer, 2007*
PFS (median, months)			
Intermediate/Poor	5.3	8.4	-
Intermediate	6.1	-	11
Poor	2.7	-	4

*Registration trial for sunitinib, MSKCC risk criteria used in this study

- ERG also noted that CABOSUN investigators suggest that the differences in PFS may reflect the inclusion of patients with poorer prognostic factors than other trials
  - ERG considers this a plausible explanation for differences, however:

	CABOSUN	CheckMate 214
Ratio of intermediate:poor risk patients	4.3:1	3.7:1

31

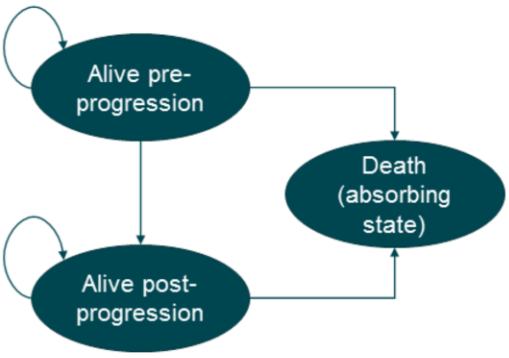
Key issues – clinical effectiveness

- Where will the technology be used in the treatment pathway?
- Is the clinical evidence generalisable to UK clinical practice?
- Which overall survival data should be used to inform the economic model?
- Is use of the indirect treatment comparison to assess clinical effectiveness valid?
  - How important are the uncertainties introduced by the differing populations in CABOSUN and COMPARZ?
  - Does the proportional hazards assumption hold for OS?
  - Is the choice of curves used to model survival appropriate?

Cost effectiveness evidence

Company submission section 3

Modelling approach and structure



- Partitioned-survival model
- Estimated proportions in each health state based on parametric survival curves fitted to clinical trial data on PFS and OS
- Based on patient-level data for comparison with sunitinib and ITC for comparison with pazopanib
- Time horizon: 20 years
- Cycle length: 1 week

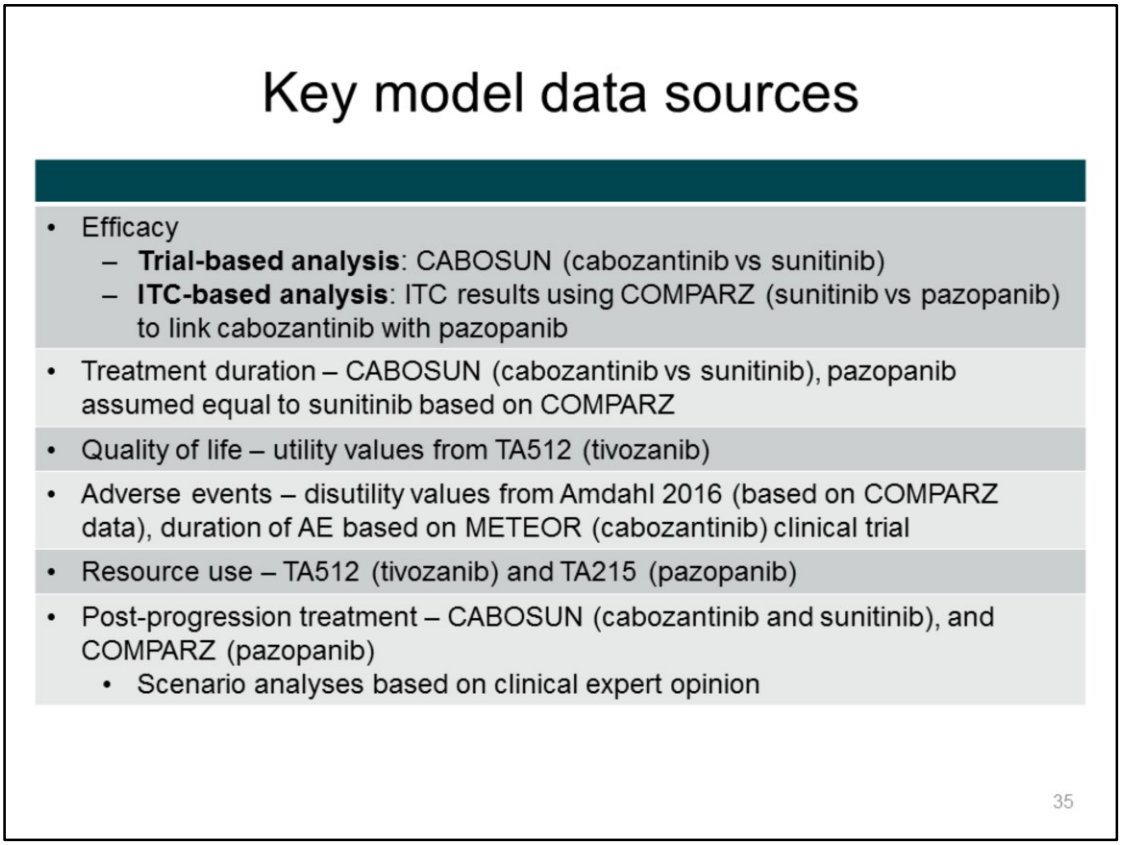
Treatment	Dosage regimen
Cabozantinib (oral)	60 mg daily, continuously administered
Sunitinib (oral)	50 mg daily for 4 weeks followed by 2 weeks without treatment
Pazopanib (oral)	800 mg daily, continuously administered

ERG comments: appropriate structure, cycle length and time horizon. Effect of shorter time horizon tested in scenario analyses – little effect on ICER with 10 year time horizon (used in previous RCC appraisals)

34

Source: company submission, p79–87

Sensitivity analysis shows that shortening time horizon to 5 years increases ICERs by ~10k



Company submission document A, p16–19 table 5 and main submission, p118, section B3.6, table 59

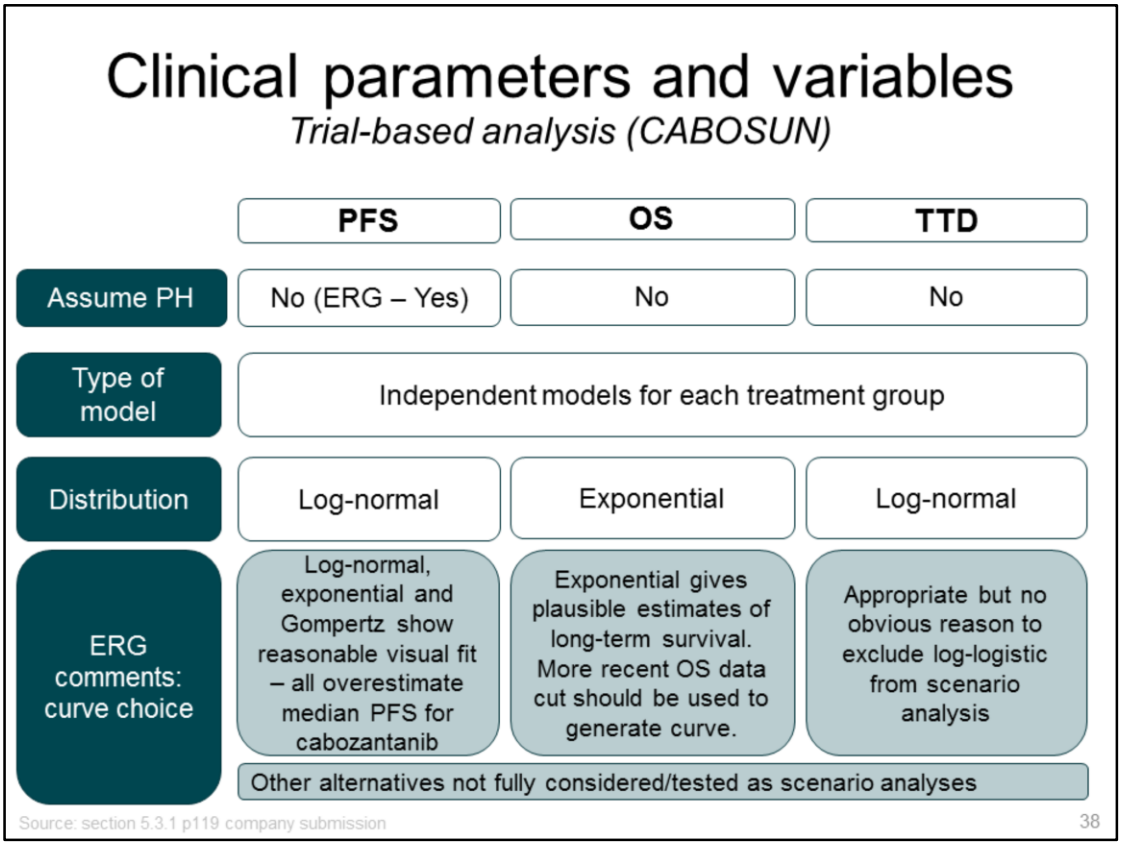
Key model assumptions

	Company assumption	ERG preferred assumptions	ERG additional comments on ERG preferred assumptions	Company and ERG's scenarios
OS curves	Jan 2017 data cut Direct comparison: Exponential Indirect comparison: FP (P1=P2=-1)	**July 2017 data cut** **Exponential** **Pazopanib is equivalent to sunitinib:**	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 COMPARZ.	**Company model:** Direct: Weibull, Gompertz, lognormal, loglogistic, gamma Indirect: Weibull, Gompertz, exponential, lognormal, loglogistic, alternative FP distributions **ERG model:** Application of alternative OS hazard ratios:HR = 0.74 (Jan 2017 analysis), no effect (HR=1),
	No adjustment for age-related mortality	**Age-related mortality**	Mortality rate can't be lower than that of general population (ONS 2014-16).	
PFS curves	Direct comparison: Lognormal Indirect comparison: FP (P1=P2=-1)	Direct comparison: Lognormal **Pazopanib is equivalent to sunitinib**	Same as in company direct base case. Lognormal gives most plausible fit.	**Company model:** Direct: exponential, Weibull, Gompertz, loglogistic, gamma Indirect: Weibull, Gompertz, exponential, lognormal, loglogistic, alternative FP distributions **ERG model:** Separate exponential, Weibull, Gompertz

Key model assumptions (continued)

	Company assumption	ERG preferred assumptions	ERG additional comments on ERG preferred assumptions	Scenarios
TTD curves	Lognormal direct comparison Pazopanib is equivalent to sunitinib	Same as company (direct comparison)	Lognormal gives the best fit, but there **is little reason to choose between other functions**, so all should be considered in scenario analyses	Exponential, Gompertz, Weibull, loglogistic, gamma
Persistence of treatment benefit	Full time horizon	**No effect 5 years after starting treatment**	Given the weakness of evidence for the OS difference, take conservative approach, with **progression and mortality hazards for cabozantinib equal to those of sunitinib after 5 years** (3 years after trial follow up).	**ERG model:** no effect 10/ 20 years after starting treatment
Subsequent treatment costs	Use of second-line treatments from trials (CABOSUN and COMPARZ)	Same as company (trial data)	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. **ERG tests 2 other scenarios** • ERG 1: **equal distribution** of NICE approved second-line drugs and • ERG 2: based on **clinical advice** on most likely distribution of second-line treatments	Company and ERG scenarios as described in previous column.

37



Although exponential conflicts with PH assumption as mentioned on previous slides ERG considers uncertainty in OS data means that OS curves shouldn't be over interpreted (ie there is uncertainty over whether PH holds).

Clinical parameters and variables

Indirect treatment comparison-based analysis

	PFS	OS
Assume PH	No	No – uncertainty due to small sample size
Type of model	Independent models for each treatment group	
Distribution	**Fractional polynomial P1=P2=-1**	**Fractional polynomial P1=P2=-1**
ERG comments: curve choice	Good fit but long-term projections seem optimistic **Lognormal** (used in ERG base case) and **loglogistic** provide best balance of fit to CABOSUN data with realistic long-term extrapolation	Small sample size and relative immaturity of OS data make it difficult to assess the fit and extrapolation of ITC curves. **Random effects exponential** and **fractional polynomial (P1=P2=-1)** both reasonable with no clear reason to choose between them, Weibull similar but has low estimates of long term survival with standard treatment.
	Other alternatives not fully considered/tested as scenario analyses	

Source: section 5.3.1 p119 company submission

ERG critique of survival inputs

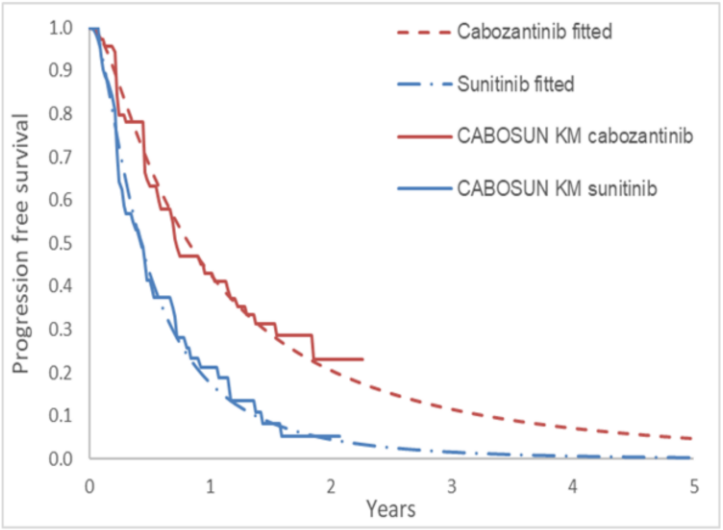
- The ERG notes some uncertainties in the use of the indirect treatment comparison (ITC) fore estimating PFS:
  - Uncertainty over the robustness of the ITC results due to differences in the trial populations
  - An overestimation of median PFS for cabozantinib compared to the CABOSUN trial result for all extrapolations
  - ERG preferred direct comparison (lognormal distribution) and assumption that pazopanib has equivalent efficacy to sunitinib to use of ITC
- ERG: Use of earlier OS data cut introduces bias in favour of cabozantinib
  - Preferred use of most recent OS data cut (July 2017) and considered Jan 2017 data cut in scenario analyses
- ERG's preferred approach to OS mdelling
  - Estimated sunitinib curve from CABOSUN data (exponential fit)
  - Applied the OS hazard ratio to generate the OS curve for cabozantinib
  - Pazopanib assumed to have same OS as sunitinib based on results from COMPARZ

40

Survival extrapolations

Trial-based analysis (CABOSUN) – Progression-free survival

PFS – lognormal distribution



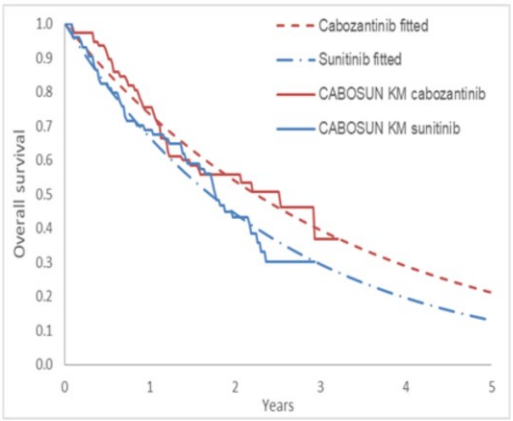
- ERG: Reasonable visual fit, although overestimates median PFS for cabozantinib – also used in ERG base case

41

Survival extrapolations

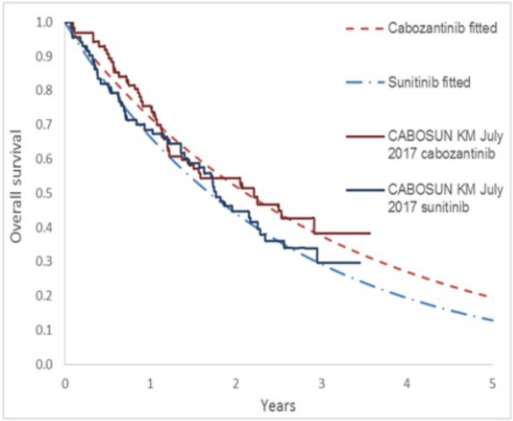
Trial-based analysis (CABOSUN) – Overall survival

Company base case
OS (Jan 2017 data cut)



Exponential curves fitted separately to both arms

ERG base case
OS (July 2017 data cut)



Exponential curves fitted to sunitinib arm
Cabozantinib curve generated using HR=0.80 from July 2017 data cut

- Proportional hazards do not hold → exponential not normally appropriate

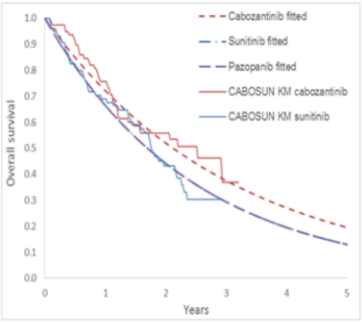
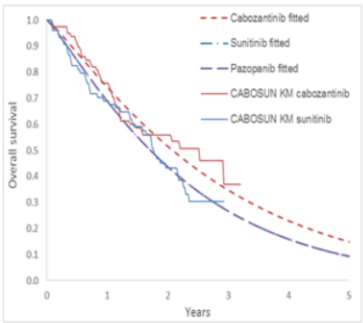
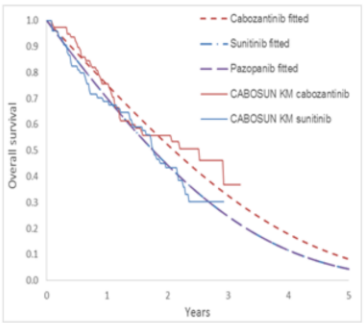
42

ERG base case assumes sunitinib and pazopanib are equivalent.

ERG notes caution in over interpreting OS so chose exponential in spite of PH not holding as gave best statistical fit

Survival extrapolations

Trial-based analysis (CABOSUN) – alternative OS distributions

	ERG base case Exponential	ERG scenario Weibull	ERG scenario Gompertz
			
Effect on ICER compared to ERG base case		• ICER vs sunitinib increases by ~£3.5k • ICER vs pazopanib increases by ~£6.5k	• ICER vs sunitinib increases by ~£5.0k • ICER vs pazopanib increases by ~£8.5k

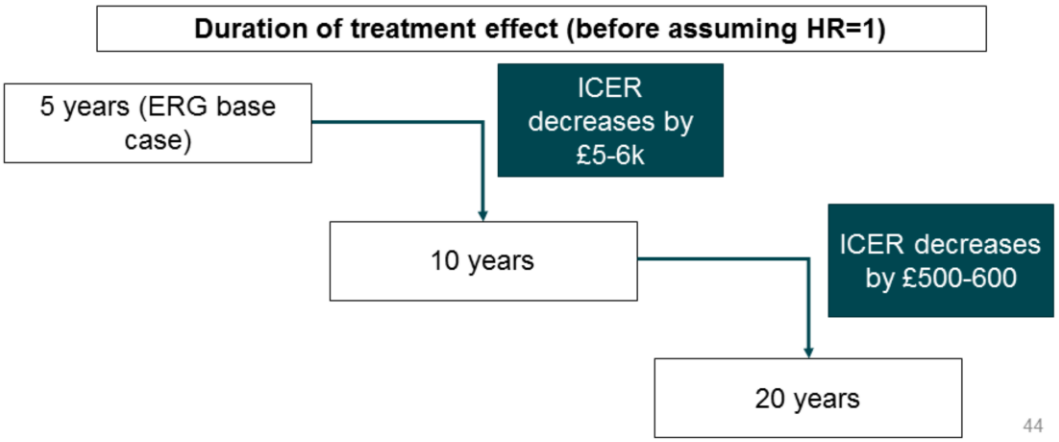
43

ERG base case assumes sunitinib and pazopanib are equivalent.

ERG notes caution in over interpreting OS so chose exponential in spite of PH not holding as gave best statistical fit

Duration of treatment effect

- Company base case assumed treatment benefit persists over entire time horizon (HR<1)
- ERG preferred assumption that treatment benefit would not persist
- ERG base case assumes equivalent efficacy for cabozantinib and sunitnib (i.e. HR=1) at 5 years with 10 and 20 years tested as scenarios



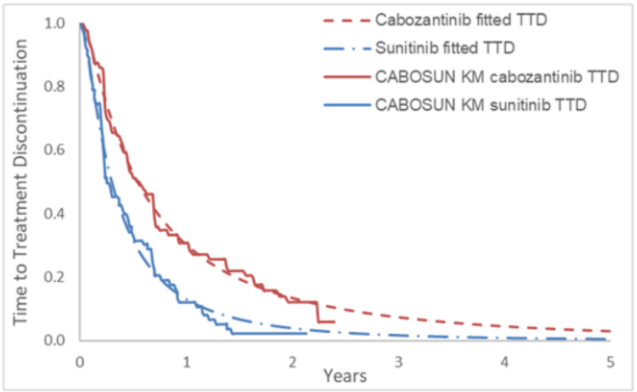
44

Time to discontinuation

Trial-based analysis (CABOSUN)

- Company and ERG base case – time to discontinuation for pazopanib assumed to be same as sunitinib as the mean treatment duration for both treatments in the COMPARZ trial was 11.5 months

Company and ERG base case – lognormal



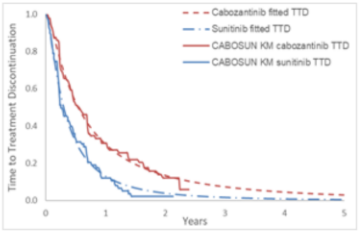
- Scenario analyses: loglogistic has no effect on ICER, other distributions (exponential, Weibull, Gompertz, gamma) result in decrease of ICER by ~£3k-£7k vs both sunitinib and pazopanib

45

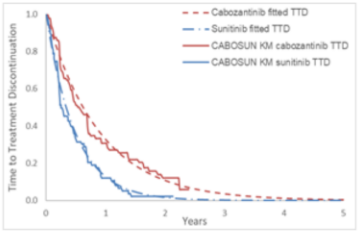
Time to discontinuation less than PFS likely due to withdrawals due to adverse events.

Time to discontinuation

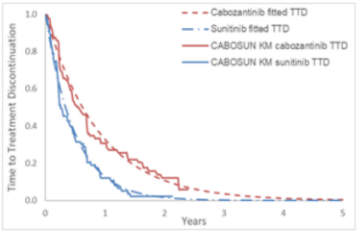
Scenarios (ERG assumptions)



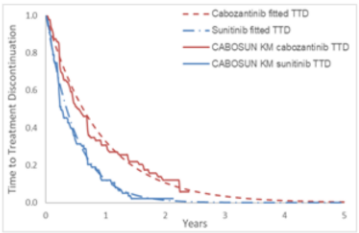
Lognormal (base case)



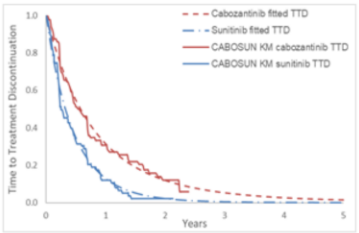
Exponential
ICER ↓ ~7k



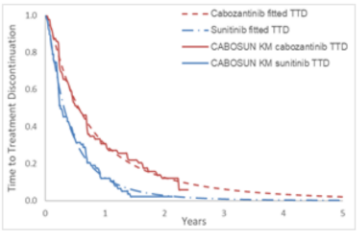
Weibull
ICER ↓ ~7k



Loglogistic
No effect on ICER



Gompertz
ICER ↓ ~3k



Gamma
ICER ↓ ~2k

46

Model inputs

Utility values

- No HRQoL data were collected in CABOSUN
- A systematic literature review (SLR) was carried out to identify relevant HRQoL data
- Clinicians recommended to company that utility values from sunitinib HTA be used as this was direct comparator in CABOSUN
- Company's base case employs utility values from TA512 (Tivozanib) as this is most recent HTA published (other values examined in scenario analyses)
- ERG agrees with company preference for utility values from TA512

	Utility value (mean)			
Source	**Tivozanib HTA**	**Sunitinib HTA**	**Pazopanib HTA**	**Swinburn 2010**
Pre-progression	0.726	0.780	0.70	0.795*
Post-progression	0.649	0.705	0.59	0.355

47

*State description: Stable disease, no adverse event

Model inputs

Disutilities due to adverse events

- Company preferred disutility values from COMPARZ trial as results could be extended to the comparator (pazopanib), where no direct evidence exists
- Company noted that this is the lowest value for adverse event-related disutilities and is a conservative estimate
- Company assumed AE duration of 4 weeks based on data from METEOR trial

State	Duration of adverse events	Number of episodes experienced per patient	QALY decrement	Source
TEAE, grade 3/4	4 weeks – assumption	1.429	-0.2044	Amdahl, 2016 (COMPARZ)

- ERG: Mean QALY loss of 0.0225 per TEAE reflects a reasonable average. Higher and lower estimates tested in scenario analyses – no significant changes to ICER
- ERG: Also assessed impact of reducing inclusion threshold for incidence from >5% to >2% and increasing duration of events to 8 weeks – no significant changes to ICER in either case

48

TEAE, treatment emergent adverse event

Source: Company submission p101–102, section 3.4, table 45

CONFIDENTIAL

Costs and resource use

Treatment	Dose regimen	PAS discount	List price	Relative dose intensity % (SE)	Mean cost per week
Cabozantinib	60 mg daily administered continuously	Confidential PAS applied	20/40/60 mg tabs x 30: £5,143	94.3 (1.5)	[redacted]
Sunitinib	50 mg daily for 2 weeks followed by 2 weeks rest	No charge for first cycle. List price thereafter	50 mg caps x 28: £3,138	87.4 (6.3)	First 6 weeks: nil Thereafter: £457.22
Pazopanib	800 mg daily administered continuously	12.5% discount on all doses	400 mg tabs x 30: £1,121	86.0 (8.6)	£393.66

- Drug cost per cycle determined based on dosage after accounting for relative dose intensity
- Flat price applied for cabozantinib doses so reduction does not impact on cost
  - Value for relative dose intensity of cabozantinib is percentage of days without dose interruption due to adverse evens
- No additional administration costs considered for oral chemotherapies
- ERG: assumptions reasonable, alternative dose intensities considered as scenarios

49

Source: Company submission, p106, section B3.5, table 48

Subsequent therapy – trial data

	Cabozantinib CABOSUN (%)	Sunitinib CABOSUN) (%)	Pazopanib COMPARZ (%)
Axitinib	23	19	6
Pazopanib	16	12	0
Sunitnib	13	13	29
Temsirolimus	9	4	6
Nivolumab	13	15	0
Everolimus	8	19	31
Sorafenib	1	3	11
Bevacizumab	0	6	7
Cabozantinib	1	6	0
Interferon	1	0	0

- Company and ERG base case both use data from CABOSUN and COMPARZ
- ERG notes that model tends to overestimate average duration and cost of second-line therapy as the same mortality rate is assumed before and after first-line treatment

50

Note: Nivolumab not in standard use when COMPARZ was carried out so use of nivo likely underestimated

Subsequent therapy – scenario analyses

	Cabozantinib (%)			Sunitinib/Pazopanib (%)		
	Company	ERG 1	ERG 2	Company	ERG 1	ERG 2
Axitinib	50	10	0	40	10	0
Nivolumab	30	30	45	30	20	30
Everolimus	10	20	0	10	20	0
Cabozantinib	0	0	0	10	20	30
Lenvatinib + everolimus	0	30	45	0	20	30
BSC	10	10	10	10	10	10

- Assumptions account for cost of subsequent therapy but not clinical effect
- Company scenario analysis conducted using clinical expert opinion elicited during TA512 (Tivozanib) and redistributing patients receiving supportive care to cabozantinib. Lenvatinib+everolimus was not a treatment option at that time
- ERG 1 reflects using only NICE recommended second-line drugs and ERG 2 clinical opinion received by ERG on likely use of second line treatments
- ICERs sensitive to choice of scenario

BSC, Best supportive care

51

CONFIDENTIAL

End-of-life criteria

Criterion	Company submission	ERG comments
Life expectancy less than 24 months	**Criterion met** IMDC validation study (Heng, 2013): median OS with sunitnib from the start of treatment • Intermediate risk: 22.5 months (18.7-25.1) • Poor risk: 7.8 months (6.5-9.7) Note: CABOSUN trial data show median OS with sunitinib of: • Intermediate/poor risk: 21.2 (16.3-27.4) • Intermediate risk: 23.5 (18.9, 28.1) • Poor risk: 6.4 (2.2, 22.4)	**Criterion not met** **Mean OS from model in both company and ERG base exceeds 24 months** for sunitinib and pazopanib (approximately 6 months without discounting in ERG analysis)
Extension to life greater than 3 months	**Criterion met** 9.3 month difference in median OS after a median follow-up of 28.9 months In company model cabozantinib is associated with a gain of greater than 3 months vs both sunitinib and pazopanib	**Criterion met** ERG analysis confirms cabozantinib is associated with a gain of greater than 3 months vs both sunitinib and pazopanib

Source: Adapted from company submission

Reference: Lancet Oncol. 2013 Feb;14(2):141-8. doi: 10.1016/S1470-2045(12)70559-om company submission, p65, section 2.13, table 26

Cost-effectiveness results

- Results confidential due to use of comparator PAS information. Please see confidential appendix accompanying this PMB

Sensitivity analysis

- Company's one-way deterministic sensitivity analyses suggest that:
  - **cost of interventions**
  - **relative dose intensity of the treatment and comparator**

  are key drivers of cost-effectiveness

- ERG: this is misleading as effectiveness parameters are not included in these analyses. Impact of these uncertainties is reflected in the probabilistic sensitivity analysis and scenario analyses

54

ERG comments

Conclusions

- ERG identified a number of uncertainties in the company's model relating to:
  - the method of fitting the OS curves
  - the persistence of treatment effects
  - health state utilities, adverse effects and costs
- As in the company' base case, sunitinib is dominated by pazopanib due to its higher cost and similar effectiveness
  - Result sensitive to some cost and resource use assumptions
  - Results generally robust, with the ICERs for cabozantinib compared with pazopanib (with and without subsequent treatment patient access schemes taken into account) remaining above £30,000 per QALY gained for all scenarios tested

55

Key issues – cost effectiveness

- Is a direct analysis of clinical effectiveness data comparing cabozantinib to sunitinib preferred to the indirect treatment comparison?
- Is the assumption that pazopanib and sunitinib have equivalent efficacy appropriate to inform the model?
- Is it reasonable to assume a persistent treatment benefit for cabozantinib?
  - If not how long should treatment benefit persist?
- Should subsequent treatment options be based on clinical trial data or clinical opinion on use in the NHS?
- Does cabozantinib meet end of life criteria?
- What is the most plausible ICER?

56

Equality considerations and innovation

- No equality considerations were identified
- The company highlighted the 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 (MET, AXL), as well as tumour angiogenesis (VEGF). By targeting MET and AXL receptors in addition to VEGFR, cabozantinib may provide additional anticancer efficacy over the more selective, existing anti-VEGFR agents.
- Cabozantinib was granted Promising Innovative Medicine (PIM) designation under the Early Access to Medicines Scheme (EAMS) in July 2016
- The company and patient groups both highlighted that cabozantinib may be more effective in the treatment of bone metastases
- ERG noted that in TA463 (cabozantinib for previously treated RCC) it was accepted that cabozantinib may have additional benefits for some patients and could be considered innovative, but cabozantinib did not represent a 'step-change' in treatment in that population

57

Authors

- **Alan Lamb**
  Technical Lead
- **Ahmed Elsada**
  Technical Adviser
- with input from the Lead Team: Stuart Williams, Nicky Welton, Nigel Westwood

58

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma [ID1208]

Document A

Company evidence submission summary for committee

Ipsen Ltd UK confirm that all information in the submission summary is an accurate summary or replication of evidence in the main submission and accompanying appendices and that wherever possible a cross reference to the original source is provided.

January 2018

File name	Version	Contains confidential information	Date
ID1208_Cabozantinib_Document A_05 Feb 18_Updated 18 Apr 18_ACIC	Final (2)	Yes	5 February 2018

Contents

Tables

Figures

Submission summary

A.1 Health condition

Renal cell carcinoma (RCC) accounts for 80% of kidney cancer cases.[1,2] Advanced RCC includes locally advanced RCC that cannot be removed by surgery, and metastatic RCC. Symptoms include fatigue, weight loss, anaemia, hypertension, fever, cachexia (wasting), neuromyopathy and amyloidosis.[3] Additional symptoms related to metastatic spread include bone pain, skeletal-related events and hypercalcaemia; lung symptoms such as airway obstruction; and venous thromboembolism.[4,5] RCC can significantly affect patients' health-related quality of life (HRQL), including physical function and psychosocial wellbeing.[6,7]

Cabozantinib is expected to be licensed for treatment of 'advanced RCC in treatment-naïve adults with intermediate or poor risk per IMDC criteria'.[8] Intermediate and poor risk patients have a poor prognosis:

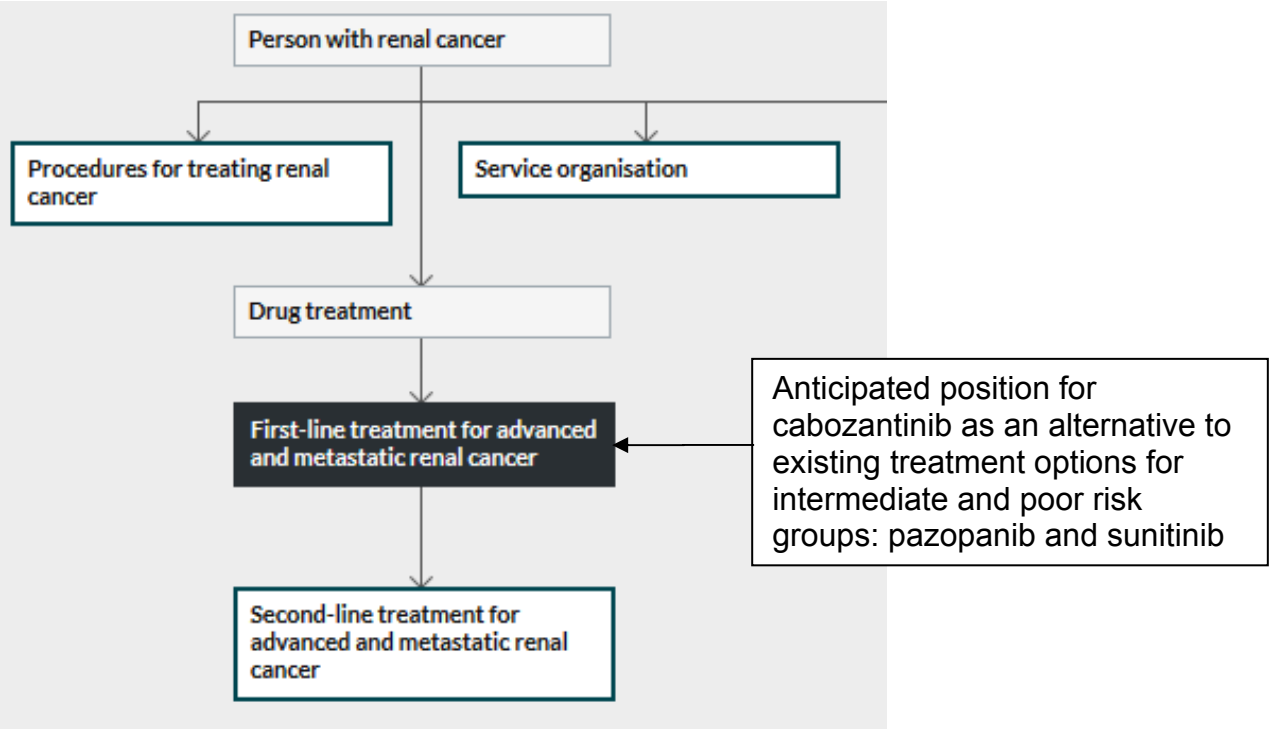
- In the International Metastatic RCC Database Consortium (IMDC) model validation study (1028 previously untreated patients receiving VEGF-targeted treatment for metastatic RCC), median overall survival (OS) from start of treatment was 43.2 months (95% CI 31.4-50.1) in the favourable, 22.5 months (18.7-25.1) in the intermediate and 7.8 months (6.5-9.7) in the poor-risk group.[9]
- In trials, median progression-free survival (PFS) in previously untreated advanced RCC in the whole population ranges from 8 to 11 months with sunitinib or pazopanib.[10-12] However, an analysis of the IMDC cohort found that median PFS with targeted agents in first line was only 5.6 months, when the population is restricted to intermediate or poor risk patients. [13]
- The 5-year relative survival rate for metastatic RCC in the UK is approximately 6%.[14]

In addition, bone metastases have a negative effect on survival in patients treated with current targeted treatments.[15]

A.2 Clinical pathway of care

Advanced RCC is incurable and is largely resistant to chemotherapy, radiotherapy and hormonal therapy. Current NICE-recommended medicines for previously untreated (first-line) advanced RCC are sunitinib[16] and pazopanib,[17] both oral VEGF-targeted agents. It is anticipated that cabozantinib will be used in accordance with its anticipated marketing authorisation for the 'treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria' and in this position cabozantinib will represent an additional treatment option alongside sunitinib and pazopanib in these intermediate or poor risk patient groups.

Figure 1 NICE pathway of care in renal cancer



Source: NICE[18]

A.3 Equality considerations

Not applicable: the use of cabozantinib is not likely to raise any equality issues.

A.4 The technology

Table 1 Technology being appraised – (B.1.2, Table 2, page 11)

UK approved name and brand name	Approved name: cabozantinib Brand name: CABOMETYX®▼
Mechanism of action	Cabozantinib is a multi-targeted therapy that inhibits multiple receptor tyrosine kinases (RTKs) implicated in tumour growth and angiogenesis, pathologic bone remodelling, drug resistance, and metastatic progression of cancer.[8] As well as inhibiting the MET (hepatocyte growth factor receptor protein) and VEGF (vascular endothelial growth factor) receptors, it also inhibits AXL and a number of other RTKs.[8] Targeting MET and AXL in addition to VEGF receptors (VEGFRs) may provide additional anticancer effects in patients with RCC compared with more selective VEGFR-inhibition strategies.[19]
Marketing authorisation/CE mark status	Application for marketing authorisation for cabozantinib for 'the treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria' was submitted to the European Medicines Agency (EMA) on 28 August 2017. CHMP Opinion received: 22 March 2018 Anticipated date of approval: May 2018 Cabozantinib (CABOMETYX) has previously received marketing authorisation for the treatment of advanced RCC in adults following prior VEGF-targeted therapy: this was granted by the EMA on 9 September 2016.[8]
Indications and any restriction(s) as described in the summary of product characteristics	The anticipated indication for cabozantinib with regard to this submission is 'treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria'.[8] Therapy with cabozantinib should be initiated by a physician experienced in the administration of anticancer medicinal products.
Method of administration and dosage	Cabozantinib as CABOMETYX® is for oral use. The recommended dose is 60 mg once daily. Treatment should continue until the patient is no longer clinically benefitting from therapy or until unacceptable toxicity occurs.[8]

	Management of suspected adverse drug reactions may require temporary treatment interruption and/or dose reduction. When dose reduction is necessary, it is recommended to reduce to 40 mg daily, and then to 20 mg daily.[8] Dose interruptions are recommended for management of CTCAE grade 3 or greater toxicities or intolerable grade 2 toxicities. Dose reductions are recommended for events that, if persistent, could become serious or intolerable.[8]
Additional tests or investigations	None. The investigations needed to establish intermediate or poor risk per IMDC criteria are carried out as part of routine clinical practice.
List price and average cost of a course of treatment	£5,143.00 for a 30 tablet pack (list price). [redacted] (average cost of an annual course of treatment with a PAS) (Budget Impact Analysis, 4.2, Table 6, page 17).
Patient access scheme (if applicable)	The simple discount scheme already in place for cabozantinib for the treatment of advanced RCC in adults following prior VEGF-targeted therapy, will be applied.

A.5 Decision problem and NICE reference case

The submission covers the technology's full anticipated marketing authorisation for this indication and is consistent with the final NICE scope and the NICE reference case.

Table 2 The decision problem – (B.1.1, Table 1, page 10)

	Final scope issued by NICE/reference case	**Decision problem addressed in the company submission**	**Rationale if different from the final NICE scope**
Population	People with untreated, intermediate or poor risk (as per International Metastatic Renal Cell Carcinoma Database Criteria), locally advanced or metastatic renal cell carcinoma	As per the scope	Not applicable
Intervention	Cabozantinib	As per the scope	Not applicable

Comparator(s)	Pazopanib Sunitinib	As per the scope	Not applicable
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment • health-related quality of life	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment	Quality of life data were not collected in the single phase II (CABOSUN) trial and data from published sources are used in the economic model.
Source of data for measurement of health-related quality of life	TIVO-1 study (trial-derived values, reported in tivozanib NICE submission[20])	As per scope	Not applicable
Source of preference data for valuation of changes in health-related quality of life	EQ-5D-3L data were obtained from the above source.	As per scope	Not applicable

A.6 Clinical effectiveness evidence

Evidence on the efficacy and safety of cabozantinib in previously untreated advanced RCC is available from one Randomised Control Trial (RCT), the Alliance A031203 CABOSUN Trial (NCT01835158).[19,21,22]

CABOSUN was an open label, phase 2, RCT comparing cabozantinib with sunitinib in patients with previously untreated advanced or metastatic RCC of poor or intermediate risk as defined by the IMDC criteria.

Both treatments were given at their approved (or anticipated approved) dosage and schedule. The open label design was chosen in order to enable appropriate dose modifications of study treatments to manage adverse events (AEs). Crossover between treatment arms was not permitted. However, as would be expected, many patients received subsequent anti-cancer therapy after discontinuing study treatment.

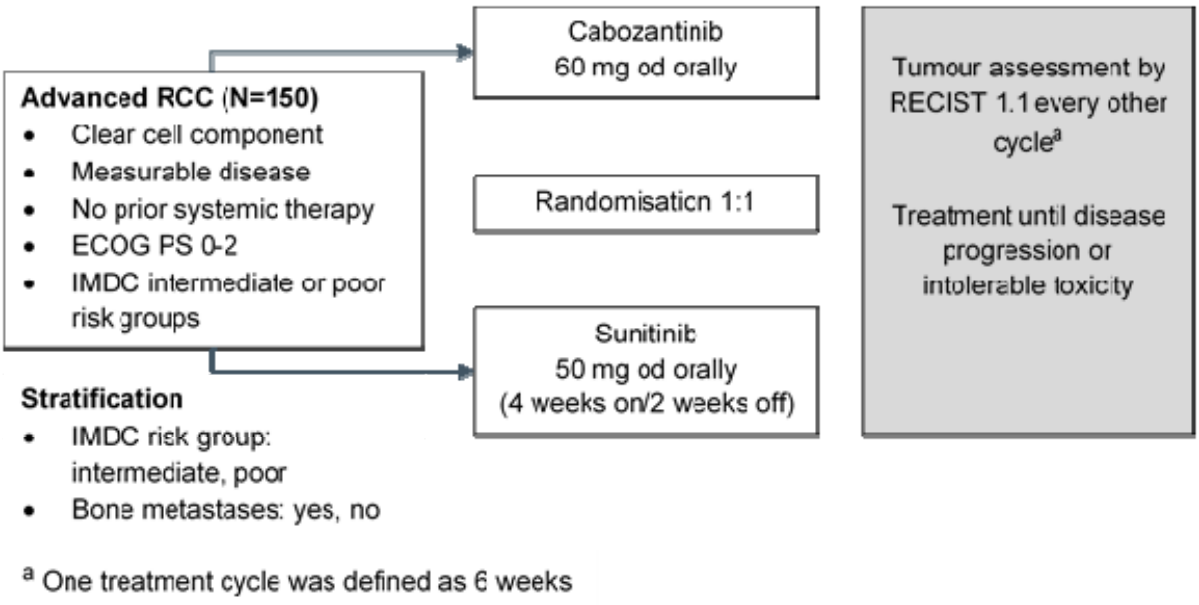
The primary outcome was PFS, defined as the interval between randomisation and first documentation of disease progression, or death from any cause. When the study was conducted, progression was investigator-assessed. For the regulatory submission, a blinded central review of the radiographic images was carried out retrospectively by an independent radiology committee (IRC), to determine progression and response. The results reported in this submission to NICE are IRC-assessed.

The secondary end points were:

- ORR, defined as proportion of subjects at the time of data cut-off with a best overall response of complete response (CR) or partial response (PR), which was confirmed by a subsequent visit ≥ 28 days later (assessment as for PFS, above)
- OS, defined as time from randomisation to death from any cause
- AEs

The study design is shown in Figure 2 with further details in Table 3.

Figure 2 Study design (B2.2, Figure 4, Page 30)



ECOG PS, Eastern Cooperative Oncology Group performance status; od, once daily; RECIST, Response Evaluation Criteria In Solid Tumors

CABOSUN was an investigator-led study carried out by the Alliance for Clinical Trials in Oncology and originally published by Choueiri et al. 2016[19]; these results used investigator assessment of progression and response, Alliance censoring rules for progression, and 1-sided p-values. Additional analyses were subsequently performed by the manufacturer for regulatory purposes, using assessment by IRC, FDA-recommended censoring rules, and two-sided p-values – the results from these analyses are included in the clinical study report (CSR). This submission is based on these analyses (as reported in the CSR[21] and published by Choueiri et al. 2017[22]). Updated OS analysis (data cut-off July 2017) as included in Choueiri et al. 2017[22] is also provided.

Table 3 Clinical effectiveness evidence (B.2.2, Table 8, page 28)

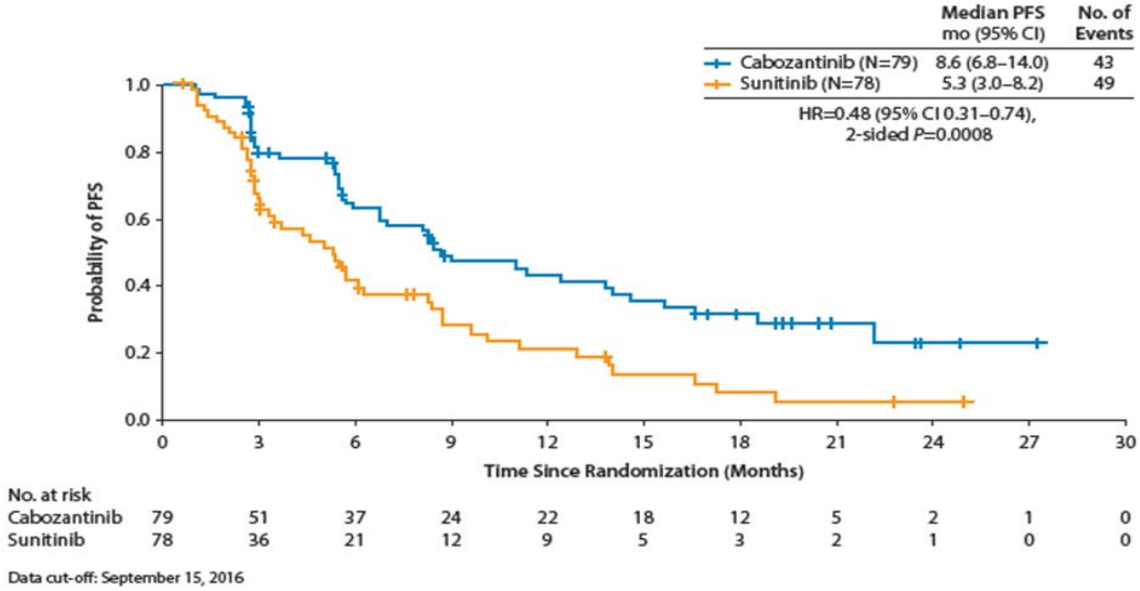
Study	CABOSUN[19,21,22]		
Study design	Open label RCT (phase 2)		
Population	Patients with untreated clear cell metastatic RCC, ECOG performance status of 0 to 2 and intermediate or poor risk per IMDC criteria		
Intervention(s)	Cabozantinib (60 mg once per day)		
Comparator(s)	Sunitinib (50 mg once per day; 4 weeks on, 2 weeks off)		
Indicate if trial supports application for marketing authorisation	Yes	**Indicate if trial used in the economic model**	Yes
Rationale for use/non-use in the model	CABOSUN was used in the model because it is the trial on which the regulatory submission is based, and it is the only trial of cabozantinib in the population stated above. The economic model was based on the regulatory analyses as published in the CSR.		
Reported outcomes specified in the decision problem	**Progression-free survival, overall survival**, objective response rate, **adverse effects of treatment** * Outcomes used in the economic model are given in bold		
All other reported outcomes	Duration of response		

A.7 Key results of the clinical effectiveness evidence

A.7.1 *Progression-free survival (primary end point)*

Cabozantinib treatment resulted in a clinically meaningful and statistically significant prolongation of PFS compared with sunitinib, by IRC assessment (8.6 vs. 5.3 months, difference 3.3 months, HR 0.48 [95% CI 0.31, 0.74], p=0.0008; Figure 3).[21,22]

Figure 3 PFS by IRC assessment (B.2.6, Figure 5, page 38)



CI, confidence interval; HR, hazard ratio; PFS, progression-free survival. Source: Clinical study report[21], Choueiri et al. 2017[22]

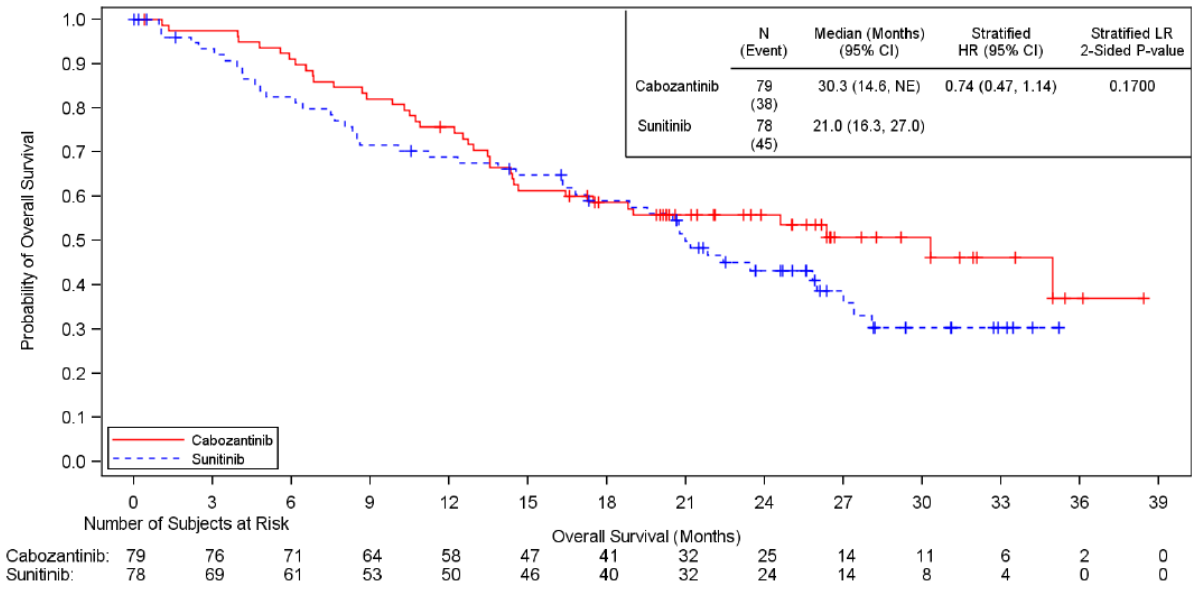
The PFS benefit obtained with cabozantinib was not diminished in patients who had bone metastases: the Hazard Ratios (HRs) in patients with and without bone metastases were 0.51 (95% CI 0.26, 0.99) and 0.50 (95% CI 0.29, 0.85), respectively.[21,22] (B.2.6, page 37)

A.7.2 *Overall survival (secondary end point)*

At the data cut-off point used for analysis of OS in the CSR (13 January 2017), and used in the economic model, median follow-up for OS was 28.9 months. The survival data were immature – only 38 deaths were recorded in the cabozantinib arm and 45 in the sunitinib arm.[21]

Cabozantinib was associated with a substantial numerical improvement in median OS compared with sunitinib. The Kaplan-Meier estimates for median OS were 30.3 months (95% CI 14.6, not estimable) 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 (Figure 4). However, the confidence intervals around the medians and HR remained wide due to the relatively low number of deaths, and the HR (0.74 in favour of cabozantinib [95% CI 0.47, 1.14; stratified 2-sided log-rank p-value = 0.1700] did not reach statistical significance.[21] (B.2.6, page 39)

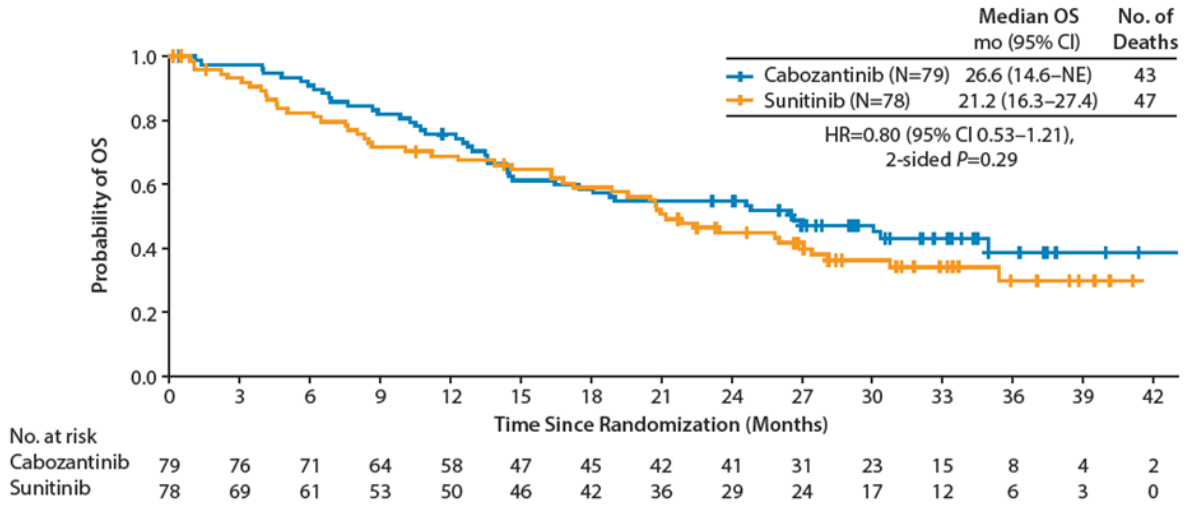
Figure 4 Overall survival, ITT population, data cut-off 13 January 2017 (B.2.6, Figure 6, page 40)



Stratification factors were IMDC risk category (intermediate, poor) and bone metastases (yes, no); CI, confidence interval; HR, hazard ratio; LR, log rank; NE, not estimable. Source: Clinical study report[21]

Updated OS results with a data cut-off of 1 July 2017[22] (Figure 5) demonstrated that median OS continued to favour cabozantinib, at 26.6 (95% CI 14.6, NE) vs. 21.2 (95% CI 16.3, 27) months for cabozantinib and sunitinib, respectively. The HR was 0.80 in favour of cabozantinib (95% CI 0.53, 1.21, 2-sided p-value = 0.29). (B.2.6, page 40)

Figure 5 Overall survival, ITT population, data cut-off 1 July 2017 (B.2.6, Figure 7, page 41)



CI, confidence interval; HR, hazard ratio; OS, overall survival. Source: Choueiri et al 2017[22]

A.7.3 *Objective response rate by IRC*

The IRC-determined ORR with cabozantinib was twice that seen with sunitinib: 20% (95% CI 12.0, 30.8) in the cabozantinib arm, compared with 9% (95% CI 3.7, 17.6) in the sunitinib arm (ITT population). All responses were partial responses. The difference in ORR between treatment arms was 11.3% (95% CI 0.4, 22.2, stratified 2-sided p-value = 0.0406).[21,22]

The disease control rate (complete response + partial response + stable disease) was markedly higher with cabozantinib than with sunitinib, at 75% and 47%, respectively.[22] (B.2.6, page 41)

A.7.4 *Adverse reactions*

The overall incidence of all-causality AEs in CABOSUN was generally similar between the cabozantinib and sunitinib arms, and the adverse events were similar to those observed with other VEGFR-TKIs in RCC.[21] The proportion of patients with grade 3 or 4 AEs was also similar between cabozantinib and sunitinib (68% and 65%, respectively).[21,22] The proportion of patients discontinuing study treatment due to AEs was similar for each treatment (21% and 22% for cabozantinib and sunitinib, respectively).[21,22] The most frequent treatment-related Grade 3/4 AEs (≥ 5% of patients) in the cabozantinib arm were hypertension (22.0%), diarrhoea (9.0%), hypophosphatemia (9.0%), palmar-plantar erythrodysaesthesia syndrome (7.7%), fatigue (5.1%), decreased appetite (5.1%), and stomatitis (5.1%). (B.2.10, page 58)

A.8 Evidence synthesis

In the absence of any direct head-to head data comparing cabozantinib with pazopanib, an indirect treatment comparison (ITC) was performed with the results used to inform OS and PFS inputs for the economic model. (B.2.9, page 45) The ITC also included a comparison of cabozantinib with sunitinib. Exponential, Weibull, Gompertz, log-logistic and log-normal distributions were included in the ITC. In addition, a fractional polynomial ITC was also applied to generate the curve parameters. The choice of base case was based on statistical fit and clinical opinion. The fractional polynomial second order model (P1=-1, P2=-1) provided the best fit to the ITC data and was therefore used in the base case for the comparisons between

cabozantinib and pazopanib. Other distributions were included in scenario analyses. (B.3.3, page 86)

The ITC results suggest that cabozantinib significantly increases PFS compared with sunitinib and pazopanib. Overall survival was longest for cabozantinib.

A.9 Key clinical issues

CABOSUN trial

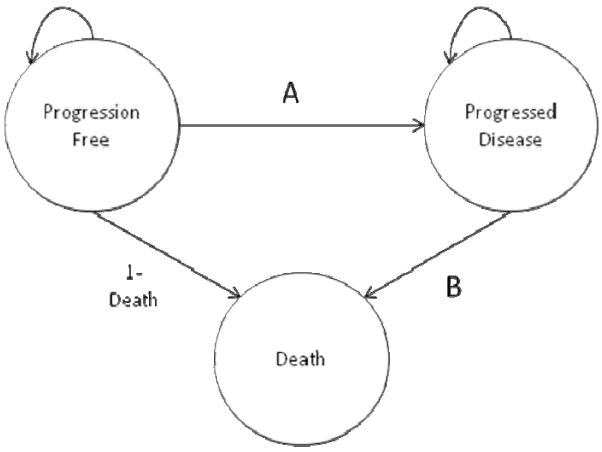
- The study was not powered to detect differences in OS. Due to the OS data from CABOSUN being immature, attributed to the relatively low number of deaths, only numerical improvements in OS for cabozantinib compared with sunitinib are reported. However, while the survival data are still immature, the latest OS results (Figure 5, above) were consistent with those previously reported, supporting the efficacy of cabozantinib (B.2.6, page 39-41).
- No crossover was allowed but, as would be expected, patients in both arms received subsequent anti-cancer therapies after discontinuing study treatment (B.2.6, Table 16, Page 44).
- The number of patients withdrawing from the study or with missing or unevaluable data was greater in the sunitinib arm than in the cabozantinib arm. The reasons for this are discussed in (B.2.13, page 64). It is unlikely that this imbalance biased the study findings against sunitinib (B.2.13, page 62).
- PFS with sunitinib in CABOSUN was lower than in some other studies, but the observed PFS is consistent with analyses from other trials and real-world data sets that examine sunitinib in a comparable population (intermediate and poor risk patients only)[13,23] (B.2.13, page 63).

ITC

- In the ITC, there were differences between the study populations in the trial comparing cabozantinib and sunitinib (CABOSUN [21]) and the trial comparing pazopanib and sunitinib (COMPARZ[10]): COMPARZ included favourable-risk patients, whereas only intermediate and poor risk patients were eligible for CABOSUN. The impact of differences in risk category could not be accounted for because patient-level data from COMPARZ were not available to access (B.2.9, page 56).

A.10 Overview of the economic analysis

Figure 6 Model diagram – (B.3.2, Figure 12, page 81)



Model structure: three-state partitioned survival model.

Comparators: Pazopanib and sunitinib.

Cycle length: one week, with half-cycle correction to account for the differences in treatment schedules.

Time horizon: 20 years (captures the life expectancy of the cohort; impact explored in sensitivity analysis).

Transition probabilities: the proportion of patients who are in each health state at any given time with sunitinib and cabozantinib, is estimated based on parametric survival curves fitted to CABOSUN clinical trial survival data over time (e.g. the proportion in the progressed disease state is calculated as the difference between the extrapolated OS and PFS curves).

For the comparison with pazopanib, PFS and OS curves were generated using ITC data based on CABOSUN and the COMPARZ study of sunitinib vs. pazopanib.[10,24]

Two analyses were performed: a CABOSUN based analysis and an indirect treatment comparison-based analysis.

A.11 Incorporating clinical evidence into the model

PFS and OS

Data from the CABOSUN study were used to inform the cost-effectiveness comparison of cabozantinib versus sunitinib. Comparisons with sunitinib and pazopanib are supported by results from the ITC (B.2.9, page 45).

The OS and PFS data from the CABOSUN study were used to calculate the proportion of patients in each treatment arm who were in each health state at any time point after starting treatment. The proportion of patients in the post-progression health state at any given time was calculated as the difference between the OS and PFS curves (B.3.3, page 86).

The assumptions of proportional hazards and accelerated failure time were tested. The proportional hazards assumption holds for the COMPARZ PFS and OS data, but not for the CABOSUN PFS and OS data. It was decided that separate fits were to be used, and that comparison of survival curves would be a more appropriate method than comparison of hazard ratios. Cabozantinib, sunitinib and pazopanib were compared using several models: Weibull, Gompertz, log-logistic, log-normal and exponential based on the fixed effect (FE) and random effect (RE) models, [25] and first and second order fractional polynomials.[26] (B.3.3, page 87-93).

Extrapolation of survival was done by applying the base case curves until the end of the time horizon. When choosing the base case survival curves clinical opinion was sought.

Time to treatment discontinuation

Data from the CABOSUN study were used in the model to inform the comparison of cabozantinib and sunitinib. For pazopanib in the absence of any data, equivalence to sunitinib was assumed. (B.3.3, page 93)

Adverse events

Treatment-related AEs were included in the model if they affected ≥ 5% of the population in any of the pivotal trials, and were judged by clinical experts to have implications for resource use.

Utility inputs

The CABOSUN trial did not collect EQ-5D or any other generic preference-based measure to estimate utilities. Other quality of life data to enable mapping to a generic measure were also not collected. Given the difficulties in combining utility estimates from different sources, including differences in trial populations and/or elicitation methods, and the lack of patient-reported data from CABOSUN, the base case analysis uses utility values derived from the TIVO-1 study as reported in the tivozanib NICE technology appraisal for the progression-free and post-progression states. For disutilities associated with AEs data from a pazopanib study were applied. Table 4 summarises the utility values used in the base case (B.3.4 Table 46, page 103)

Table 4 Base case utility values

State	Utility Value	95% CI	Source
Progression-free	0.726 (0.011)	0.705, 0.748	Tivozanib Appraisal TA10123[20]
Post-progression	0.649 (0.019)	0.612, 0.686	
TEAEs Grade ¾	-0.2044 (0.0682)	-0.0707,0.3381	Amdahl 2016[27]

A.12 Key model assumptions and inputs

Table 5 Key model assumptions and inputs (B.3.6, Table 59, page 118)

Model input and cross reference	Source/assumption	Justification
OS and PFS	CABOSUN study [21] (cabozantinib vs. sunitinib) and COMPARZ study [10](pazopanib vs. sunitinib)	The relative efficacy for cabozantinib vs pazopanib in the model is based on an ITC, which assumes that there are no significant imbalances in treatment effect modifiers between different studies. There are differences in risk category between the study populations (COMPARZ included favourable risk patients whereas CABOSUN did not), but it was not possible to re-run the ITC for particular subgroups due to lack of data.
Quality of life	Quality of life is dependent on disease progression status and toxicity of treatments	Given the difficulties in combining utility estimates from different sources, including differences in trial populations and/or elicitation methods, and the lack of patient-reported data from CABOSUN, the base case analysis uses utility values derived from the TIVO-1 study as reported in the recent tivozanib NICE technology appraisal (see Table 4 above).[20] The utilities were estimated on the full TIVO-1 trial population, not on the treatment-naïve population.
Treatment duration	Treatment duration was characterised by log-normal curve for cabozantinib and sunitinib. For pazopanib, no TTD	No TTD Kaplan-Meier data were identified for pazopanib. Using PFS as a proxy for TTD is likely to result in over-estimation of TTD. This method is used to avoid over-estimation of treatment duration in the pazopanib

	data were found from the literature search. It was assumed that pazopanib TTD is equal to sunitinib TTD.[28]	arm. The assumption is based on the COMPARZ study, where mean treatment durations for both sunitinib and pazopanib was 11.5 months.[28]
Abbreviations: ITC, indirect treatment comparison; TTD, time to treatment discontinuation; OS, overall survival; PFS, progression-free survival		

A.13 Base-case ICER (deterministic)

Table 6 and Table 7 summarise the results for the base case (with the PAS) based on the CABOSUN study and the ITC results

Table 6 Base-case results (deterministic): pairwise analysis of cabozantinib versus sunitinib (from CABOSUN study) – (B.3.7, Table 60, page 123)

Drug	Total costs	Total QALYs	Total life-years	ΔCosts	ΔQALYs	ΔLife years	ICER vs baseline	ICER
Sunitinib	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	15,170	0.401	0.657	37,793	37,793
Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years								

Table 7 Base-case results (deterministic): pairwise analysis of cabozantinib versus comparators (based on the ITC results) (B.3.7, Table 61, page 123)

Drug	Total costs	Total QALYs	Total life-years	ΔCosts	ΔQALYs	ΔLife years	ICER vs baseline	ICER
Pazopanib	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Sunitinib	[redacted]	[redacted]	[redacted]	7,561	-0.021	-0.035	Dominated by pazopanib	Dominated by pazopanib
Cabozantinib	[redacted]	[redacted]	[redacted]	23,526	0.486	0.799	48,451	31,538
Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years								

A.14 Probabilistic sensitivity analysis

Table 8 Mean probabilistic base case results – (B.3.8, Tables 62-64, page 125)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental. costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Incremental ICER (£/QALY)
Cabozantinib vs. sunitinib (CABOSUN-based)								
Sunitinib	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	14,722	0.392	0.645	37,592	37,592
Cabozantinib versus comparators (based on the ITC results)								
Pazopanib	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Sunitinib	[redacted]	[redacted]	[redacted]	7,293	(0.055)	(0.115)	Dominated by Pazopanib	Dominated by Pazopanib
Cabozantinib	[redacted]	[redacted]	[redacted]	16,673	0.551	0.943	48,289	30,239

ITC-based results for cabozantinib vs. sunitinib and cabozantinib vs. pazopanib are included in B.3.8 Tables 63 and 64

Figure 7 Scatter plot of probabilistic results, cabozantinib vs sunitinib – (B.3.8, Figure 17, page 126)

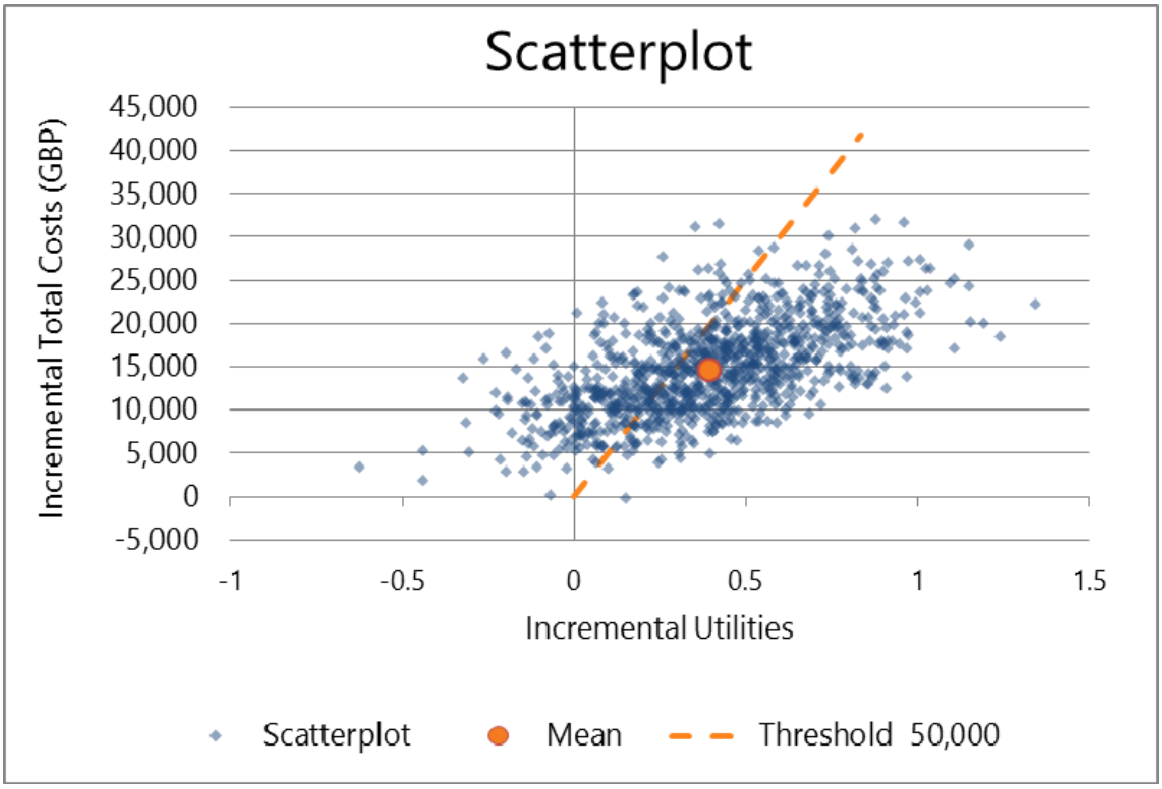
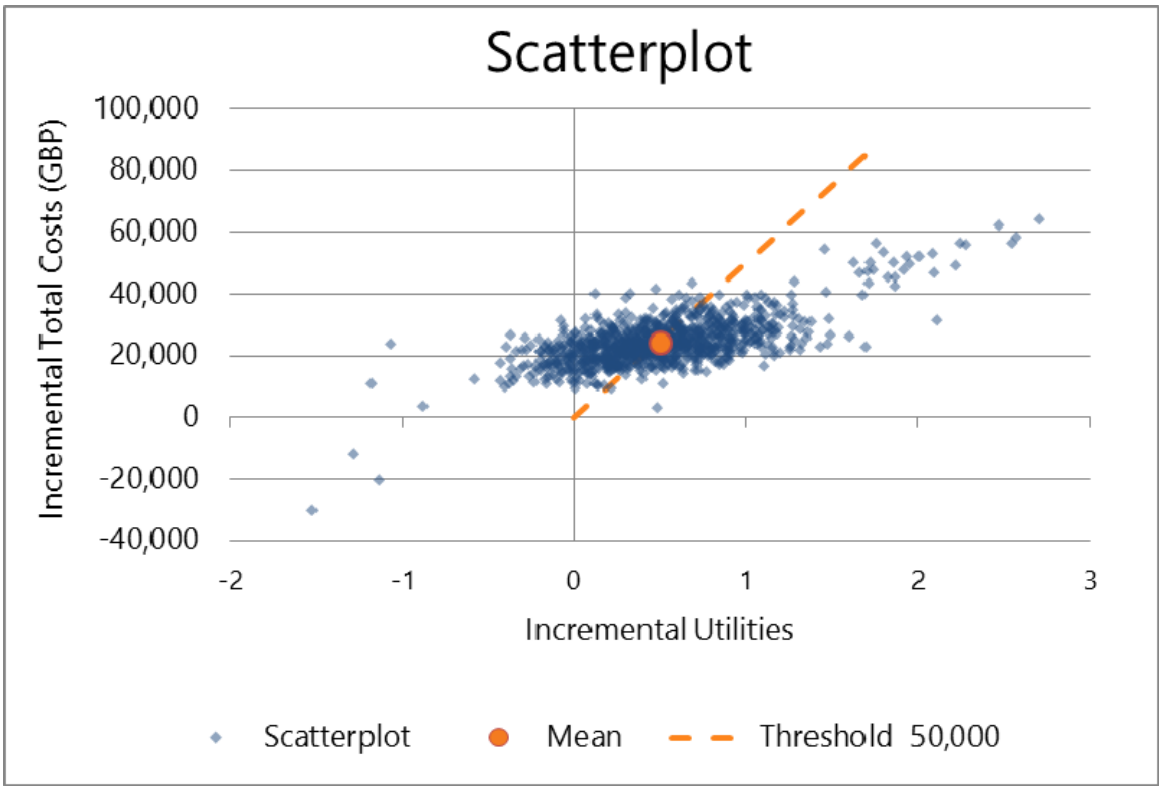


Figure 8 Scatter plot of probabilistic results, cabozantinib vs pazopanib – (B.3.8, Figure 19, page 127)



A.15 Key sensitivity and scenario analyses

Figure 9 Tornado diagram, cabozantinib vs sunitinib – (B.3.8, Figure 23, page 133)

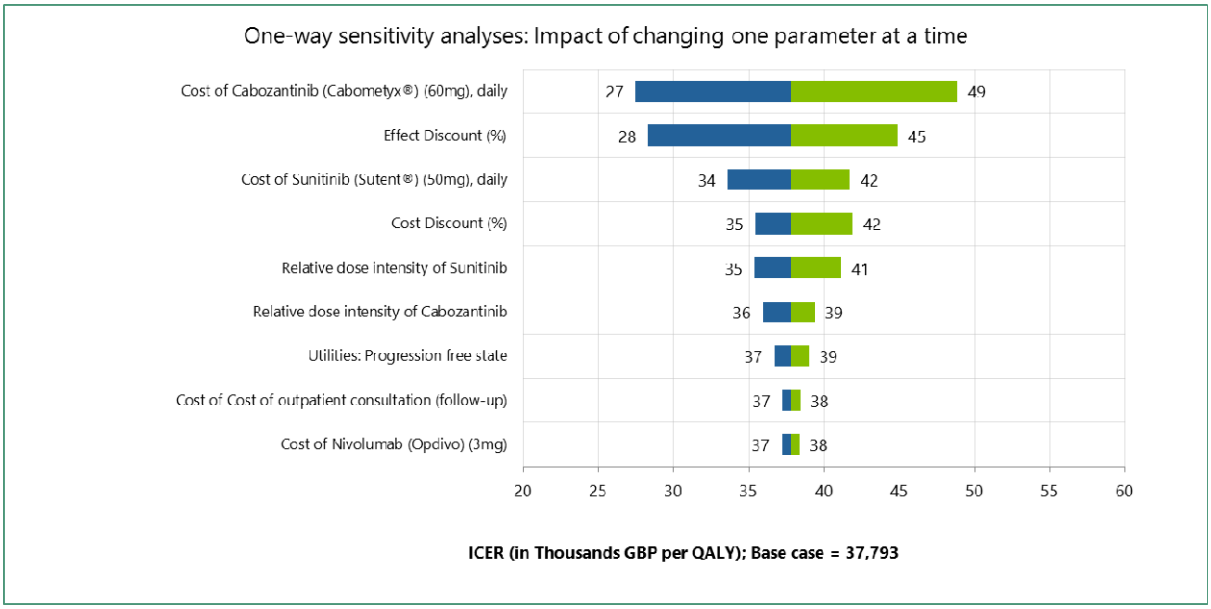


Figure 10 Tornado diagram, cabozantinib vs pazopanib – (B.3.8, Figure 24, page 134)

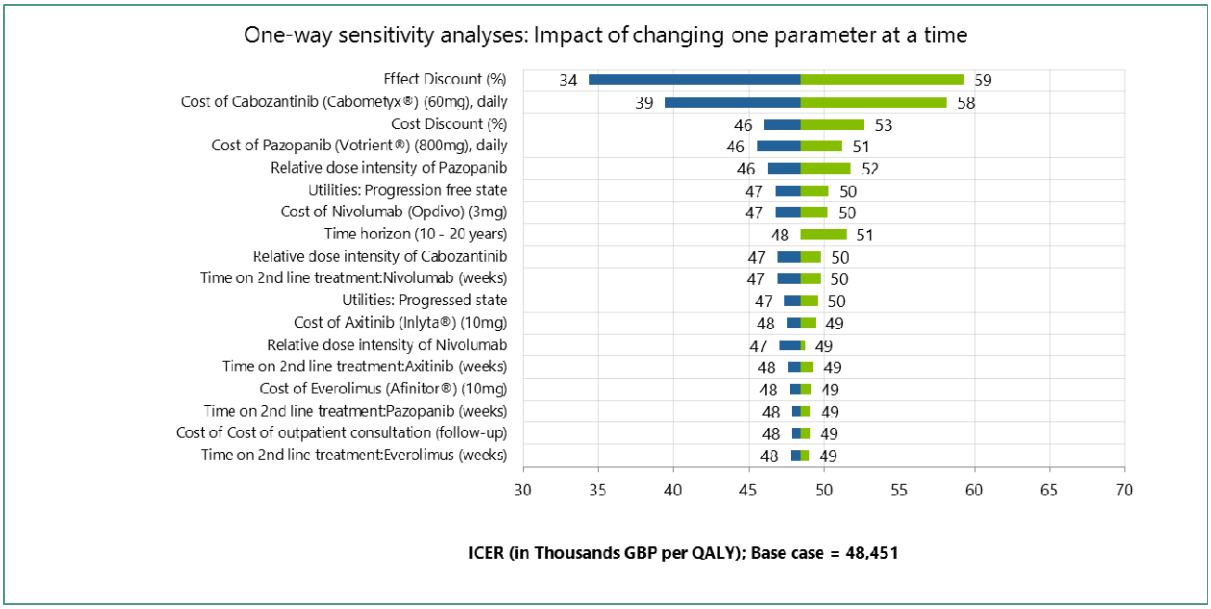


Table 9 Key scenario analyses – versus sunitinib (B.3.8, Table 66, page 136)

Scenario and cross reference	Scenario detail	Brief rationale	Base case ICER	Scenario ICER	Impact on base-case ICER
Base case			**37,793**		
OS curve choice	Weibull	The ICER is sensitive to changes in OS curve choice, because it affects how long people live i.e. total life-years and QALYs gained.		46,660	+8,867
	Gompertz			34,895	-2,898
	2nd order model (P1=-0.5, P2=0)			55,541	+17,747
	2nd order model (P1= -1, P2=0)			38,735	+942
TTD curve choice	Exponential	The ICER is sensitive to changes in treatment duration assumptions because it impacts the total primary intervention cost.		31,458	-6,335
	Weibull			31,470	-6,323
	Gompertz			34,285	-3,509
	Gamma			35,345	-2,449
Utility values	Swinburn[29]	The ICER is sensitive to changes in utility value because it impacts the total QALYs estimated for each treatment.		31,196	-6,598
	Pazopanib NICE STA[30]			38,727	+ 934
	Sunitinib NICE STA[16]			35,985	-1,808

Table 10 Key scenario analyses – versus pazopanib (B.3.8, Table 67, Page 137)

Scenario and cross reference	Scenario detail	Brief rationale	Base case ICER	Scenario ICER	Impact on base-case ICER
Base case			**48,451**		
OS curve choice	2nd order model (P1=-0.5, P2=0)	The ICER is sensitive to changes in OS curve choice, because it affects how long people live i.e. total life-years and QALYs gained.		81,252	+32,802
	2nd order model (P1= -1, P2=0)			58,566	+10,115
TTD curve choice	Exponential	The ICER is sensitive to changes in treatment duration assumptions because it impacts the total primary intervention cost.		42,967	-5,484
	Weibull			42,908	-5,542
	Gompertz			45,440	-3,011
	Gamma			46,107	-2,344
Utility values	Swinburn[29]	The ICER is sensitive to changes in utility value because it impacts the total QALYs estimated for each treatment.		35,870	-12,580
	Pazopanib NICE STA[30]			48,946	+495
	Sunitinib NICE STA[16]			46,684	-1,767

A.16 Innovation (B.2.12, page 62)

The superior efficacy results of cabozantinib seen both in CABOSUN and in the METEOR trial of previously treated advanced RCC[31,32] may 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 (MET, AXL), as well as tumour angiogenesis (VEGF).[33] By targeting MET and AXL receptors in addition to VEGFR, cabozantinib may provide additional anticancer efficacy over the more selective, existing anti-VEGFR agents (B.2.12, page 62).

Cabozantinib was granted Promising Innovative Medicine (PIM) designation under the Early Access to Medicines Scheme (EAMS) in July 2016.

A.17 End-of-life criteria

The end-of-life criteria are considered to apply to cabozantinib treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria (see Table 11).

Table 11 End-of-life criteria – (B.2.13, Table 28, page 65)

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	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.[9]
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	In the CABOSUN trial,[21] 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[10]). In the economic modelling, cabozantinib is associated with a gain of 0.80 life years (9.6 months) compared with pazopanib.
Abbreviations: IMDC, OS, overall survival: PFS, progression-free survival; RCC, renal cell carcinoma	

A.18 Budget impact

Table 12 summarises the results of the budget impact described in Document D: Budget Impact Analysis Submission.

Table 12 Budget impact – (Budget Impact Analysis Submission, 7.1, Table 9 page 20, Table 10 page 21)

	Company estimate	Cross reference
Number of people in England who would have treatment	182	Budget impact document, Section 3.3 page 9; Section 5.1 Table 8, page 19
Average treatment cost per person	Average cost per patient in Year 1 as estimated from the cost-effectiveness analysis for cabozantinib = £[redacted]	Budget impact document, Section 4.2, Table 6, page 17
Estimated annual budget impact on the NHS in England	Net budget impact in Year 1= £[redacted] million (with PAS) £[redacted] million (list price)	Budget impact document, Section 7.1, Tables 9, 10, pages 20-21

A.19 Interpretation and conclusions of the evidence

Cabozantinib represents a clinically and cost-effective treatment option for treatment-naïve adults with advanced RCC with intermediate or poor risk per IMDC criteria.

Cabozantinib delivers clinically meaningful improvements in OS and PFS, while maintaining a manageable toxicity profile. Compared with sunitinib in previously untreated patients with advanced RCC, and as demonstrated in the CABOSUN trial, cabozantinib was associated with a:

- Statistically and clinically significant prolongation of IRC-assessed PFS compared with sunitinib (8.6 vs 5.3 months, HR 0.48 (95% CI 0.31, 0.74; 2-sided p-value = 0.0008 corresponding to a 52% reduction in the risk of disease progression or death)[21,22] (B.2.6, page 36).
- A numerical improvement in OS compared with sunitinib: median OS was 30.3 months in the cabozantinib arm vs. 21.0 months in the sunitinib arm (HR 0.74 [95% CI 0.47, 1.14]; 2-sided p-value = 0.1700)[21] (B.2.6, page 39).
- An improvement in ORR: The IRC-determined ORR with cabozantinib was twice that seen with sunitinib[21,22] The disease control rate (complete response + partial response + stable disease) was markedly higher with cabozantinib than with sunitinib, at 75% and 47%, respectively[22] (B.2.6, page 41).

In addition, ITC results suggest that cabozantinib significantly increases PFS compared with sunitinib and pazopanib. Overall survival was longest for cabozantinib (B.2.9, page 57).

The safety profile of cabozantinib is similar to that of sunitinib and other established TKIs. In general adverse effects from cabozantinib can be managed with supportive care and dose reductions[8] (B.2.10, page 57).

In treatment naive advanced RCC patients treatment with cabozantinib was more costly but also more effective in terms of LYs and QALYs gained than treatment with sunitinib or pazopanib. Specifically, cabozantinib yielded an overall ICER of £37,793/QALY compared to sunitinib, the current standard of care. Cabozantinib also extended life by 0.66 Life Years (LYs) and provided 0.401 incremental QALYs gained compared to sunitinib. Compared to pazopanib, the overall ICER was £48,451/QALY. Cabozantinib also extended life by 0.80 LYs and provided 0.49

incremental QALYs gained compared to pazopanib. The analysis was driven by the difference in PFS and OS between cabozantinib and sunitinib, as well as the difference in treatment costs (B.3.11, page 141).

The key strength of this analysis was that it was based on evidence from an ITC comparing parametric survival curves, rather than HRs. This avoids the issue of violating the proportional hazards assumption. Several methods of indirect comparison were carried out (fixed effect model, random effect model and fractional polynomials), and best statistical fit that also made clinical sense was chosen for the base case. In addition, resource use and cost inputs were populated using data reflecting UK clinical practice. Finally, the model concept, structure and inputs were reviewed by oncologists actively treating RCC in the UK, thereby ensuring that the model assumptions were clinically relevant to the UK setting and that a comprehensive array of costs was accounted for. Resource use and costs were validated to ensure that they were justifiable on the basis of existing data and clinical opinion and were subjected to sensitivity analysis (B.3.11, page 141).

A.20 References

1. Escudier B, Porta C, Schmidinger M, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2016;27(Suppl 5):v58-v68.
2. Lopez-Beltran A, Scarpelli M, Montironi R, Kirkali Z. 2004 WHO classification of the renal tumors of the adults. *Eur Urol.* 2006;49(5):798-805.
3. Ljungberg B, Hanbury DC, Kuczyk MA, et al. Renal cell carcinoma guideline. *Eur Urol.* 2007;51(6):1502-1510.
4. Kim DY, Wood CG, Karam JA. Treating the two extremes in renal cell carcinoma: management of small renal masses and cytoreductive nephrectomy in metastatic disease. *Am Soc Clin Oncol Educ Book.* 2014:e214-221.
5. National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology: Kidney Cancer, version 2.2018.* 2017.
6. Gupta K, Miller JD, Li JZ, Russell MW, Charbonneau C. Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): a literature review. *Cancer Treat Rev.* 2008;34(3):193-205.
7. Cella D. Beyond traditional outcomes: improving quality of life in patients with renal cell carcinoma. *Oncologist.* 2011;16 Suppl 2(Suppl 2):23-31.
8. Ipsen Pharma. *Cabometyx: Summary of Product Characteristics (Draft) (Commercial in Confidence).* 2017.
9. Heng DY, Xie W, Regan MM, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. *Lancet Oncol.* 2013;14(2):141-148.
10. Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med.* 2013;369(8):722-731.
11. Motzer RJ, Hutson TE, Tomczak P, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med.* 2007;356(2):115-124.
12. Sternberg CN, Davis ID, Mardiak J, et al. Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. *J Clin Oncol.* 2010;28(6):1061-1068.
13. Ko JJ, Choueiri TK, Rini BI, et al. First-, second-, third-line therapy for mRCC: benchmarks for trial design from the IMDC. *Br J Cancer.* 2014;110(8):1917-1922.
14. Cancer Research UK. Kidney cancer statistics. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer. Accessed January 2018.
15. McKay RR, Kroeger N, Xie W, et al. Impact of bone and liver metastases on patients with renal cell carcinoma treated with targeted therapy. *Eur Urol.* 2014;65(3):577-584.
16. National Institute for Health and Care Excellence. TA169. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma. 25 March 2009; https://www.nice.org.uk/guidance/ta169. Accessed January 2018.
17. National Institute for Health and Care Excellence. TA215. Pazopanib for the first-line treatment of advanced renal cell carcinoma. Last updated 1 August 2013; https://www.nice.org.uk/guidance/ta215. Accessed January 2018.
18. National Institute for Health and Care Excellence. Renal cancer overview. https://pathways.nice.org.uk/pathways/renal-cancer. Accessed January 2018.

19. Choueiri TK, Halabi S, Sanford BL, et al. Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Patients With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial. Published online on 14 November 2016. *J Clin Oncol.* 2017;35(6):591-597.
20. National Institute for Health and Care Excellence. Single technology appraisal. Tivozanib for treating renal cell carcinoma [ID591]. Committee papers 2017. https://www.nice.org.uk/guidance/gid-ta10123/documents/committee-papers. Accessed January 2018.
21. Exelixis Inc. *Clinical Study Report: A031203 Randomized Phase II Study Comparing Cabozantinib (NSC #761968 and IND #116059) with Commercially Supplied Sunitinib in Subjects with Previously Untreated Locally Advanced or Metastatic Renal Cell Carcinoma. Report dated 31 July 2017. (Commercial in Confidence).* 2017.
22. Choueiri T, Hessel C, Halbi S, et al. Progression-free survival by independent review and overall survival update for the Alliance A031203 CABOSUN trial of cabozantinib vs sunitinib in metastatic renal cell carcinoma. European Society for Medical Oncology Congress; September, 2017; Madrid, Spain.
23. Rini BI, Hutson TE, Figlin RA, et al. Sunitinib in patients with metastatic renal cell carcinoma: Clinical outcome according to IMDC risk group. *J Clin Oncol.* 2017;35(15,suppl):abstract 4584.
24. Motzer RJ, Hutson TE, McCann L, Deen K, Choueiri TK. Overall survival in renal-cell carcinoma with pazopanib versus sunitinib. *N Engl J Med.* 2014;370(18):1769-1770.
25. Ouwens MJ, Philips Z, Jansen JP. Network meta-analysis of parametric survival curves. *Res Synth Methods.* 2010;1(3-4):258-271.
26. Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Med Res Methodol.* 2011;11:61.
27. Amdahl J, Diaz J, Park J, Nakhaipour HR, Delea TE. Cost-effectiveness of pazopanib compared with sunitinib in metastatic renal cell carcinoma in Canada. *Curr Oncol.* 2016;23(4):e340-354.
28. The GlaxoSmithKline group of companies. *Study VEG108844, a study of pazopanib versus sunitinib in the treatment of subjects with locally advanced and/or metastatic renal cell carcinoma: overall survival update. Report dated 13 Mar 2014. (GlaxoSmithKline document number 2013N180282_00).*
29. 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):1091-1096.
30. National Institute for Health and Care Excellence. Single technology appraisal (STA). Pazopanib (Votrient®) for the first-line treatment of patients with advanced renal cell carcinoma (RCC). Manufacturer submission. 16 April 2010. https://www.nice.org.uk/guidance/ta215/documents/renal-cell-carcinoma-first-line-metastatic-pazopanib-manufacturer-submission-submission2. Accessed January 2018.
31. Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med.* 2015;373(19):1814-1823.
32. Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. *Lancet Oncol.* 2016;17(7):917-927.
33. Stukalin I, Alimohamed N, Heng DY. Contemporary Treatment of Metastatic Renal Cell Carcinoma. *Oncol Rev.* 2016;10(1):295.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma [ID1208]

Document B
Company evidence submission

January 2018

File name	Version	Contains confidential information	Date
ID1208_Cabozantinib_Document B Company Evidence Submission_05 Feb 18_Updated_18 Apr 18_ACIC	Final (2)	Yes/no	5 February 2018

Contents

Tables

Figures

List of abbreviations

AE	Adverse event
AIC	Akaike's Information Criterion
AICC	Akaike's Information Criterion (corrected)
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATA	Adequate tumour assessment
BIC	Bayesian Information Criterion
BSC	Best supportive care
CANS$	Canadian dollars
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CMH	Cochran–Mantel–Haenszel
CR	Complete response
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse events
DARE	Database of Abstracts of Reviews of Effect
DIC	Deviance information criteria
DRS	Disease related symptoms
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
EAMS	Early Access to Medicines Scheme
EAU	European Association of Urology
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EPAR	European Public Assessment Report
ESMO	European Society of Medical Oncology
FDA	US Food and Drug Administration
FE	Fixed effects
FKSI	Functional Assessment of Cancer Therapy Kidney Symptom Index
GBP	Great British Pound
HR	Hazard ratio
HRQL	Health related quality of life

HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
ICUR	Incremental cost utility ratio
IFN	Interferon
IL2	Interleukin 2
IMDC	International Metastatic RCC Database Consortium
INV	Investigator
IRC	Independent radiology committee
ITC	Indirect treatment comparison
ITT	Intent-to-treat
IV	Intravenous
LY	Life year
LYG	Life year gained
MCMC	Markov Chain Monte Carlo
MET	Hepatocyte growth factor receptor protein
mRCC	Metastatic renal cell carcinoma
MSKCC	Memorial Sloan-Kettering Cancer Center
mTOR	Mammalian target of rapamycin
NA	Not applicable
NCCN	National Comprehensive Cancer Network
NE	Not estimable
NHSEED	NHS Economic Evaluation Database
NR	Not reported
OPEN	Oncology Patient Enrollment Network
ORR	Objective response rate
OS	Overall survival
PAS	Patient Access Scheme
PD	Progressed disease
PD-1	Programmed death 1
PenTAG	Peninsula Technology Assessment Group
PF	Progression free
PFS	Progression-free survival
PH	Proportional hazard
PICO	Population, Interventions, Comparators, Outcomes

PITT	Primary endpoint intent-to-treat
PIM	Promising innovative medicine
PPES	Palmar-plantar erythrodysaesthesia syndrome
PPS	Post-progression survival
PR	Partial response
PS	Performance status
PSA	Probabilistic sensitivity analysis
QALY	Quality adjusted life year
Q-TWiST	Quality-adjusted time without symptoms of disease or toxicity of treatment
RCC	Renal cell carcinoma
RCT	Randomised controlled trial
RE	Random effects
RECIST	Response Evaluation Criteria in Solid Tumors
REL	Time after progression or relapse
RPSFT	Rank preserving structural failure time
RTK	Receptor tyrosine kinase
SAE	Serious adverse event
SIGN	Scottish Intercollegiate Guideline Network
SmPC	Summary of Product Characteristics
SRE	Skeletal related events
STA	Single technology appraisal
TEAE	Treatment emergent adverse event
TKI	Tyrosine kinase inhibitor
TNM	Tumour, Node, Metastasis
TOX	Toxicity
TTD	Time to treatment discontinuation
TRAE	Treatment related adverse event
TWiST	Time without symptoms of disease or toxicity of treatment
ULN	Upper limit of normal
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

The submission covers the technology's anticipated full marketing authorisation for this indication.

Table 1 The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with untreated, intermediate or poor risk (as per International Metastatic Renal Cell Carcinoma Database Criteria), locally advanced or metastatic renal cell carcinoma	As per the scope	Not applicable
Intervention	Cabozantinib	As per the scope	Not applicable
Comparator(s)	Pazopanib Sunitinib	As per the scope	Not applicable
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment • health-related quality of life.	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment	Quality of life data were not collected in the single phase II (CABOSUN) trial and data from published sources are used in the economic model.

B.1.2 Description of the technology being appraised

The draft summary of product characteristics (SmPC) has been submitted as part of the reference pack. The European Public Assessment Report (EPAR) is not yet available.

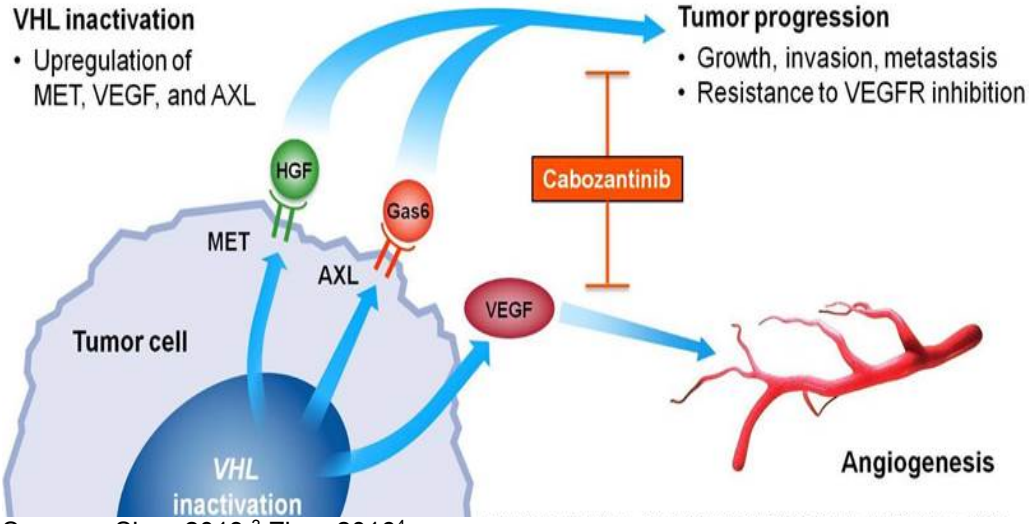
Table 2 Technology being appraised

UK approved name and brand name	Approved name: cabozantinib Brand name: CABOMETYX®▼
Mechanism of action	Cabozantinib is a multi-targeted therapy that inhibits multiple receptor tyrosine kinases (RTKs) implicated in tumour growth and angiogenesis, pathologic bone remodelling, drug resistance, and metastatic progression of cancer.[1] As well as inhibiting the MET (hepatocyte growth factor receptor protein) and VEGF (vascular endothelial growth factor) receptors, it also inhibits AXL and a number of other RTKs.[1] Targeting MET and AXL in addition to VEGF receptors (VEGFRs) may provide additional anticancer effects in patients with RCC compared with more selective VEGFR-inhibition strategies.[2] By targeting additional pathways to the VEGF pathway, cabozantinib provides a multi-targeted approach to the treatment of RCC. The mechanism of action of cabozantinib is shown in Figure 1.
Marketing authorisation/CE mark status	Application for marketing authorisation for cabozantinib for 'the treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria' was submitted to the European Medicines Agency (EMA) on 28 August 2017. CHMP opinion received: 22 March 2018 Anticipated date of approval: May 2018 Cabozantinib (CABOMETYX) has previously received marketing authorisation for the treatment of advanced RCC in adults following prior VEGF-targeted therapy: this was granted by the EMA on 9 September 2016.[1]
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The anticipated indication for cabozantinib with regard to this submission is 'treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria'.[1] Therapy with cabozantinib should be initiated by a physician experienced in the administration of anticancer medicinal products.

Method of administration and dosage	Cabozantinib as CABOMETYX is for oral use. The recommended dose is 60 mg once daily. Treatment should continue until the patient is no longer clinically benefitting from therapy or until unacceptable toxicity occurs.[1] Management of suspected adverse drug reactions may require temporary treatment interruption and/or dose reduction. When dose reduction is necessary, it is recommended to reduce to 40 mg daily, and then to 20 mg daily.[1] Dose interruptions are recommended for management of CTCAE grade 3 or greater toxicities or intolerable grade 2 toxicities. Dose reductions are recommended for events that, if persistent, could become serious or intolerable.[1]
Additional tests or investigations	None. The investigations needed to establish intermediate or poor risk per IMDC criteria are carried out as part of routine clinical practice.
List price and average cost of a course of treatment	£5,143.00 for a 30 tablet pack (list price). [redacted] (average cost of an annual course of treatment with a PAS) (Economic model).
Patient access scheme (if applicable)	The simple discount scheme already in place for cabozantinib for the treatment of advanced RCC in adults following prior VEGF-targeted therapy will be applied.

Abbreviations: CHMP, Committee for Medicinal Products for Human Use; CTCAE, Common Terminology Criteria for Adverse Events; IMDC, International Metastatic RCC Database Consortium; RCC, renal cell carcinoma

Figure 1 Cabozantinib mechanism of action



Sources: Shen 2013,[3] Zhou 2016[4]

B.1.3 Health condition and position of the technology in the treatment pathway

Summary

- Renal cell carcinoma (RCC) is a cancer that originates in the epithelium of the proximal renal tubules of the kidney and accounts for 80% of all kidney cancer cases.[5,6]
- Advanced RCC includes both locally advanced RCC that cannot be removed by surgery, and metastatic RCC.
- Advanced RCC can be associated with significant morbidity, and can significantly affect both physical and psychosocial aspects of quality of life.[7,8]
- Patients with advanced RCC are classified into risk groups, which predict survival and influence management.
- Advanced RCC is incurable: the goals of treatment are to extend life and delay disease progression, while relieving physical symptoms and maintaining function.[9]
- Previously untreated intermediate and poor risk patients (by IMDC criteria) have a poor prognosis:
  - median progression-free survival (PFS) with current agents in this population is only 5.6 months.[10]
  - median overall survival (OS) from start of first treatment is 22.5 months for intermediate risk and 7.8 months for poor risk patients.[11]
- In addition, bone metastases have a negative effect on survival in patients treated with currently available targeted treatments.[12]
- There is a need for treatments that can extend progression-free and overall survival in patients with intermediate and poor risk advanced RCC, including those with bone metastases.

Overview of renal cell carcinoma

Renal cell carcinoma (RCC) is a cancer that originates in the epithelium of the proximal renal tubules of the kidney. It accounts for approximately 80% of all kidney cancer cases.[5,6] There are several subtypes of RCC; clear cell RCC is the most common, accounting for 75% of cases.[13]

Clinical presentation

RCC is often asymptomatic in the early stages, leading to delayed diagnosis and resulting in many patients presenting with advanced disease.[14] However, with improved diagnostic imaging, incidental diagnosis of early-stage RCC is increasing, as discussed below under 'Epidemiology'.

Symptoms of advanced RCC include fatigue, weight loss, anaemia, and paraneoplastic syndromes involving hypertension, fever, cachexia (wasting), neuromyopathy and amyloidosis.[15] Additional symptoms related to the metastatic spread of the disease include bone pain, skeletal-related events and hypercalcaemia; lung symptoms such as airway obstruction; and venous thromboembolism.[14,16]

Effects of RCC on health-related quality of life

The symptoms associated with advanced disease and the generally poor prognosis for patients with advanced RCC can significantly affect all domains of patients' health-related quality of life (HRQL), including physical function and psychosocial wellbeing.[7,8] HRQL has been shown to worsen when patients with advanced disease experience disease progression.[17-20]

Staging and prognosis

Survival is dependent on the stage of the disease and on established prognostic factors (see below). RCC is staged using the 'tumour, node, metastasis' (TNM) system.[5] Advanced RCC includes both locally advanced RCC that cannot be removed by surgery, and metastatic RCC.

Patients with advanced RCC are also classified into risk groups, which predict survival and influence management. The International Metastatic RCC Database Consortium (IMDC) risk stratification model, also known as the Heng model,[11,21] is an update of the previously used Memorial Sloan Kettering Cancer Center (MSKCC) model.[22,23] The IMDC model was used in several recent trials in untreated patients,[24,25] including the CABOSUN trial of cabozantinib.[2] Its predictive value has also been validated in previously treated patients.[26] The anticipated licensed indication for cabozantinib in previously untreated advanced RCC refers to IMDC risk

groups. In the IMDC classification, patients are assessed for the presence of six risk factors:

- Karnofsky performance status (PS) <80%
- Haemoglobin <lower limit of normal
- Time from diagnosis to treatment of <1 year
- Corrected calcium above the upper limit of normal
- Platelets greater than the upper limit of normal
- Neutrophils greater than the upper limit of normal

The risk groups are: favourable (no factors), intermediate (one or two factors), and poor (more than three factors).[11] Approximately 80% of all metastatic RCC patients are in the intermediate and poor risk groups.[11] These patients have a poor prognosis:

- In the IMDC validation study (1028 previously untreated patients receiving VEGF-targeted treatment for metastatic RCC), median OS from the start of treatment was 43.2 months (95% CI 31.4-50·1) in the favourable risk group, 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.[11]
- In clinical trials, median PFS in patients with previously untreated advanced RCC ranges from 8 to 11 months with sunitinib or pazopanib for the entire patient population. [27-29] However, an analysis of the IMDC cohort found that median PFS with targeted agents in the first line setting was only 5.6 months[10] when the population is restricted to intermediate or poor risk patients.
- The 5-year relative survival rate for stage IV (i.e. metastatic) RCC in the UK is approximately 6%.[30]

Epidemiology

There were 9,023 new cases of kidney cancer (ICD-10 C64 malignant neoplasm of kidney, except renal pelvis) in England in 2015, equating to an age-standardised rate of 24.3 per 100,000 in males and 12.3 per 100,000 in females.[31] It is the UK's seventh most common cancer, accounting for 3% of incident cases.[30] The incidence of kidney cancer in the UK is projected to rise by 26% between 2014 and 2035, an

estimated annual increase of 1.2%.[30] This rise is partly due to increased incidental detection of early-stage tumours.[5]

There were 3,319 deaths from kidney cancer in England in 2015,[31] and kidney cancer accounted for approximately 3% of all cancer-related UK deaths.[30]

Kidney cancer is more common in males, with a male:female ratio of 17:10.[30] The risk of kidney cancer increases with age, and half of all cases in the UK are diagnosed in patients aged 70 or over.[30] With an ageing population and increasing prevalence of risk factors, the burden of advanced RCC is predicted to increase.[8]

Smoking, obesity and hypertension are the most important known risk factors for RCC.[5,30] In the UK, an estimated 42% of kidney cancers are linked to lifestyle factors including smoking (24%), and overweight and obesity (24%).[30]

Advanced RCC accounts for a substantial proportion of RCC cases:

- Approximately 38% of kidney cancer cases in England are diagnosed when their RCC is advanced or metastatic,[32] and 25-31% of patients present with metastases at diagnosis.[30]

Population estimates for England

The incidence of kidney cancer (ICD-10 code C64) in England in 2015 was 9,023 cases, of whom 8,927 were ≥ 15 years of age.[31] It is estimated that 1,728 new adult patients were eligible for first-line systemic treatment for intermediate or poor risk RCC in 2017, based on the assumptions that:

- the proportion of kidney cancer patients with RCC is 80%,[5,6]
- the proportion of people with RCC who present with advanced disease is 38%,[32]
- the proportion of advanced RCC patients classified as intermediate or poor risk by the IMDC criteria is 81.51%,[11]
- 75% of patients with advanced RCC are eligible for first-line systemic therapy,[33]
- the UK population is growing at 0.8% per year.

A full explanation of this estimate is given in the Budget Impact Analysis Submission Document, Section 3.3.

Clinical pathway of care

Current treatment

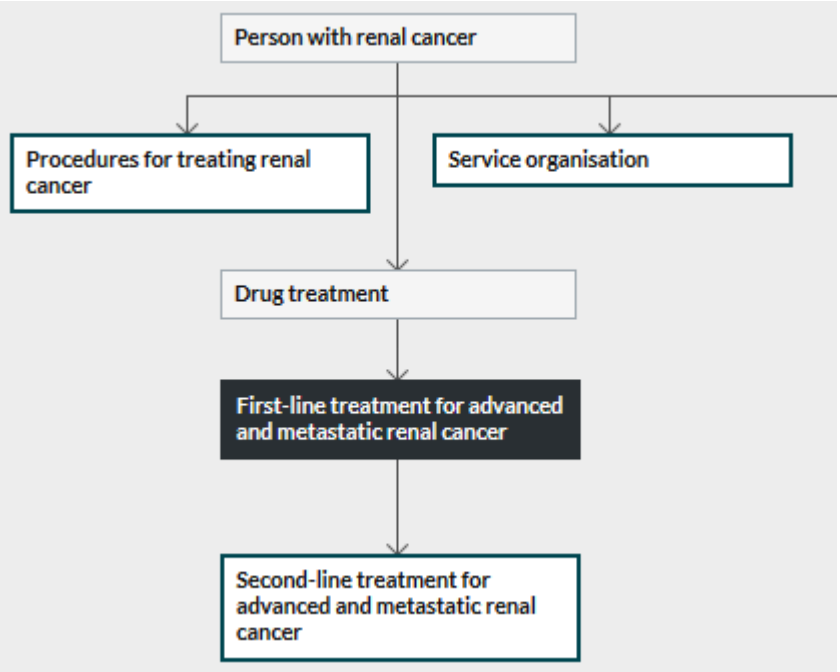
Advanced RCC is incurable, and the goals of treatment are to extend life and delay disease progression while relieving physical symptoms and maintaining function.[9]

Targeted therapies are the mainstay of treatment. Advanced RCC is largely resistant to chemotherapy, radiotherapy and hormonal therapy. Some patients receive surgery to reduce the size of the tumour or to remove metastases; this may be in addition to drug treatment.[5] The NICE pathway for RCC[34] is shown in Figure 2. Current NICE-recommended medicines for previously untreated (first-line) advanced RCC are sunitinib and pazopanib, both oral VEGF-targeted agents:

- Sunitinib is recommended as a first-line treatment option in patients who are suitable for treatment and have an ECOG performance status of 0 or 1 (TA169).[19]
- Pazopanib is recommended as a first-line treatment option in patients who have not received prior cytokine therapy and have an ECOG performance status of 0 or 1 (TA215).[35]

Bevacizumab, sorafenib and temsirolimus are not recommended (TA178).[9]

Figure 2 NICE pathway of care in renal cancer



Source: NICE[34]

There are no UK-specific clinical guidelines for the treatment of RCC. Current clinical practice in England and Wales reflects the following guidelines, whilst taking account of those medicines recommended by NICE:

- European Society of Medical Oncology (ESMO) Renal Cell Carcinoma: Clinical Practice Guidelines for diagnosis, treatment and follow-up[5]
- Updated European Association of Urology (EAU) Guidelines: Recommendations for the Treatment of First-line Metastatic Clear Cell Renal Cancer [36]
- National Comprehensive Cancer Network (NCCN) clinical practice guidelines in oncology, kidney cancer.[16]

The updated EAU guidelines and the NCCN guidelines both recommend cabozantinib as a treatment option in previously untreated IMDC intermediate and poor risk RCC.[16,36] ESMO has not yet issued guidance on the position of cabozantinib in previously untreated RCC.

Issues relating to current treatments

Current first-line treatments are associated with modest PFS in intermediate and poor risk patients.[27-29] Median PFS in patients with intermediate or poor risk per IMDC criteria is only 5.6 months.[10] OS ranges from 7.8 months in the poor risk group

to 22.5 months in the intermediate risk group.[11] Patients with bone metastases currently have a particularly poor prognosis, and have poorer outcomes with currently available targeted therapies compared with patients without bone metastases.[12,37] There is a clear unmet need for new treatments which provide clinicians and patients with an additional treatment option and can improve the prognosis for patients with intermediate and poor risk advanced RCC, as well as extend both PFS and OS in patients with and without bone metastases.

Place of cabozantinib in therapy

It is anticipated that cabozantinib will be used in accordance with its marketing authorisation, which is expected to be 'treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria'. It will represent an additional treatment option in these patients, alongside sunitinib and pazopanib in these intermediate or poor risk patient groups.

No change in current management arrangements or infrastructure will be required. The tests required to assign patients to IMDC risk groups are already carried out as part of routine clinical practice. As an oral, once-daily treatment, cabozantinib is easy to administer and offers convenience for both patients and clinicians, as it can be taken at home without the need for hospital administration, with any dose modifications managed remotely.

Note: cabozantinib is already recommended by NICE within its marketing authorisation for use in previously treated advanced RCC (NICE TA 463, August 2017).[38]

B.1.4 Equality considerations

We do not anticipate that the use of cabozantinib will be associated with any equality issues.

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

See appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised.

Identification and selection of studies

A systematic literature review was designed to identify studies on cabozantinib in first-line treatment of advanced RCC and other possible comparators including pazopanib, sunitinib, bevacizumab, interferon (IFN)-alfa, interleukin (IL)-2, sorafenib, temsirolimus, and tivozanib. The review was conducted from a global perspective and consequently included additional comparator treatments not specified in the NICE scope. A systematic literature review for 7 of 9 drugs in the indication of interest has previously been undertaken by the manufacturer of pazopanib within a submission to NICE.[18] This existing search was combined with our supplemental searches.

The manufacturer's NICE submission document for the appraisal of pazopanib included a systematic review of pazopanib, bevacizumab+IFN-alfa, IFN-alfa, IL-2, sunitinib, sorafenib and temsirolimus for the period 1980-2009. Accordingly, this report was used to identify references for this timeframe. A supplemental search was conducted for the aforementioned treatments to fill the gap for the period 2009-2017. As cabozantinib and tivozanib were also identified as first line treatments, but were not included in the pazopanib submission document, the literature search for publications on these treatments was conducted for the whole period 1980-2017.[18] For the purposes of this submission, only publications related to cabozantinib, sunitinib and pazopanib were included in the final selection.

The literature search was performed in MEDLINE®, Embase and Cochrane Library. The search in medical bibliographic databases was supplemented by a search in the clinicaltrials.gov study registry for studies providing results. During the literature search, the reference lists of identified systematic reviews, health technology assessment (HTA) submission documents, meta-analyses and indirect treatment

comparisons (ITC) were reviewed in order to obtain additional references. The search strategy is reported in detail in Appendix D 1.1.

Table 3 Data sources and date of searches

Database/ Date of search	Sources
Bibliographic Databases Date of search: 28 June 2017	MEDLINE® (Ovid MEDLINE® Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE® Daily and Ovid MEDLINE® <1946 to Present>) Embase Cochrane library, including: Cochrane Central Register of Controlled Trials Cochrane Reviews Database of Abstracts of Reviews of Effect (DARE) HTA Database NHS Economic Evaluation Database (NHSEED)
Clinical Trials Registers Date of search: 28 June 2017	Clinicaltrials.gov

Study selection

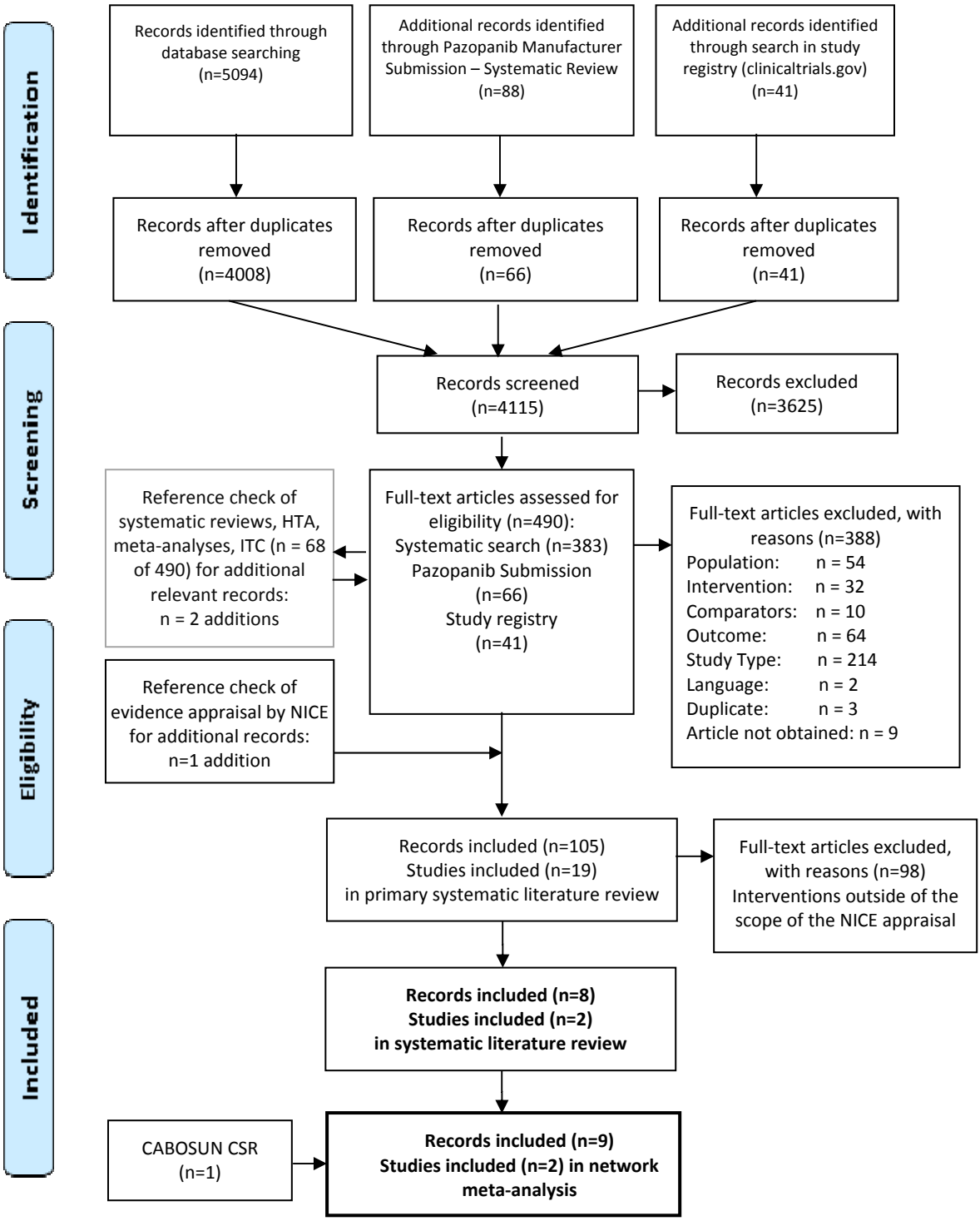
Each identified record was assessed for eligibility against predefined inclusion and exclusion criteria. The inclusion and exclusion criteria were described using the PICO(+) framework (patients, intervention, comparators, outcomes and other criteria) and are presented in Table 4. Please note that, while the global literature review considered a wide range of interventions and comparators, we have included only cabozantinib, sunitinib and pazopanib, in line with the scope for this NICE single technology appraisal.

Table 4 Inclusion and exclusion criteria applied during abstract and full-text screening

Category	Inclusion criteria	Exclusion criteria
Population	Adult patients with advanced or metastatic renal cell carcinoma, previously untreated	- Patients <18 years of age - Healthy subjects - Animal studies
Intervention	The following interventions in the first-line setting: - Cabozantinib (Cabometyx®) [monotherapy]* - Sunitinib (Sutent®) [monotherapy]* - Pazopanib (Votrient®) [monotherapy]* - Interferon alfa [monotherapy] - Interleukin-2 [monotherapy] - Sorafenib (Nexavar®) [monotherapy] - Bevacizumab (Avastin®) + Interferon alfa - Temsirolimus (Torisel®) [monotherapy] - Tivozanib (Fotivda) [monotherapy]	Combination therapies of listed interventions
Comparator	- Any from the intervention list** - Placebo - Best supportive care (BSC)	
Outcome	- Overall Survival (OS) - Progression-free survival (PFS) - Treatment discontinuation - Adverse events (AE) (incidence of AE, SAE, grade 3/4)	
Study design	- Randomised controlled trial (RCT) - Letters reporting an RCT - Systematic reviews, meta-analyses, HTA [for screening of bibliographies only, thereafter excluded]	
Language	English, French, German, Italian, Spanish	
Country	No restrictions	
Search time	1980 onwards	
Notes: *comparator treatments in the scope of this NICE appraisal ** only treatments in the scope of the NICE appraisal		

Copies of potentially relevant full papers were obtained and further selection was undertaken based on full text review. Double independent record selection was undertaken during the screening of titles/abstract as well as full texts, and discrepancies were resolved after discussion between reviewers or by a third reviewer. A PRISMA flow chart detailing the number of studies included and excluded at each stage of the review is shown in Figure 3.

Figure 3 Study flow chart



The systematic literature search in bibliographic databases retrieved in total 5,094 citations. After duplication removal 4,008 references remained for the selection process. In addition, 88 and 41 records were identified through the pazopanib manufacturer's submission document[18] and Clinicaltrials.gov study registry, respectively, and after removal of duplicates 66 and 41 references remained for the selection process, respectively. In total, 4,115 abstracts were screened and 3,625 were excluded as found not relevant.

Of the 490 full text articles, 388 were excluded. Two records were excluded due to language restriction.[39,40] Both records were in Chinese language with an English abstract. The abstracts indicated that both are systematic reviews with meta-analysis. Nine records were excluded as they could not be obtained as full text. All are conference abstracts and are listed in Appendix D. Further reasons for exclusion are listed in Appendix D. Among the excluded articles, 68 systematic reviews, HTA reports, meta-analyses and ITC were reviewed for reference checking and two additional publications were identified.[41,42] One record was a conference abstract that was not captured by the Scottish Intercollegiate Guidelines Network (SIGN) filter for RCTs used in the electronic search. The second record was a study of interferon-alfa versus interleukin-2 published in the time frame 1980-2009, which was not identified in the manufacturer's systematic review conducted for the pazopanib NICE submission. Furthermore, one additional record was identified through reviewing a relevant evidence appraisal by NICE in the indication.[43]

In total, 19 studies (105 publications) investigating the efficacy and safety of the available first line treatment options: cabozantinib, pazopanib, sunitinib, bevacizumab, IFN-alfa, IL-2, sorafenib, temsirolimus, and tivozanib were identified. Only studies comparing treatment options relevant to the scope for this appraisal (sunitinib, pazopanib or cabozantinib) were retained from the systematic literature review. Three publications referring to one study (CABOSUN) were identified as providing evidence for the clinical effectiveness of cabozantinib in untreated RCC (Table 17). For the CABOSUN study, the clinical study report (CSR) and patient-level data were also available, corresponding to a final data cut off 13 January 2017 for OS and 15 September 2016 for PFS[44] (see Section B2.2. for an explanation of the data cuts available). The full reference list of included and excluded records is

provided in Appendix D. The systematic review was also used to identify studies for potential inclusion in an indirect treatment comparison; this part of the review is described in Section B.2.9.

Table 5 Publications providing evidence on cabozantinib in untreated RCC

Study	Full reference	Main publication (main/abstract/additional)
Choueiri_2016[2]	Choueiri, T. K., Halabi, S., Sanford, B. L., et al. Cabozantinib versus sunitinib as initial targeted therapy for patients with metastatic renal cell carcinoma of poor or intermediate risk: The alliance A031203 CABOSUN trial. Journal of Clinical Oncology. 2017. 35(6): p. 591-597. Published online 14 November 2016	Main publication Final analysis
Choueiri_2016b[45]	Choueiri, T. K., Halabi, S., Sanford, B., et al. PR CABOzantinib versus SUNitinib (CABOSUN) as initial targeted therapy for patients with metastatic renal cell carcinoma (mRCC) of poor and intermediate risk groups: Results from ALLIANCE A031203 trial. Annals of Oncology. Conference: 41st European Society for Medical Oncology Congress, ESMO. 2016. 27(no pagination):	Abstract Refers to Choueiri_2016[2]
Chopra_2016[46]	Chopra, M. Annual Congress of the European Society for Medical Oncology (ESMO): Copenhagen, Denmark; 7-11 October 2016. Targeted Oncology. 2016. 11(6): p. 705-709.	Abstract Refers to Choueiri_2016[2]

B.2.2 List of relevant clinical effectiveness evidence

The CABOSUN trial

Evidence on the efficacy and safety of cabozantinib in previously untreated advanced RCC is available from one RCT, the Alliance A031203 CABOSUN Trial (NCT01835158), referred to henceforth as CABOSUN. In this trial, cabozantinib was compared with the standard of care, sunitinib, in IMDC intermediate and poor risk patients, stratified by IMDC risk group and presence of bone metastases.

CABOSUN was an investigator-led study carried out by the Alliance for Clinical Trials in Oncology, supported by grants from the National Institutes of Health, and by

Exelixis (the manufacturer of cabozantinib, who provided the drug). It was not designed as a registration trial, but in view of the encouraging results, the manufacturer and its partners decided to pursue regulatory submission based on the study.[44]

Data cut-off points and censoring

Details of the data cut-off points and censoring are provided below.

Table 6 Data cut off points and outcomes analysed

Date	Outcomes analysed	Source
11 April 2016	PFS and objective response rate (ORR)	Choueiri et al 2016[2]
15 September 2016	PFS and ORR	CSR,[44] and Choueiri et al 2017[47]
13 January 2017	OS*	CSR[44]
1 July 2017	OS*	Choueiri et al 2017[47]

*Exploratory analysis. Abbreviations: OS, overall survival; PFS, progression-free survival

11 April 2016 cut-off

The trial results were initially published by the Alliance (Choueiri et al, 2016[2]). These results used a cut-off of 11 April 2016, investigator assessment of progression and response, Alliance censoring rules for progression (under which missing or inadequate tumour assessments or use of systemic non-protocol anticancer therapy were not reasons for censoring), and 1-sided p-values. This analysis was event-driven, triggered when 123 events were observed.

15 September 2016 cut-off

Additional analyses were subsequently performed by the manufacturer for regulatory purposes, using a data cut-off of 15 September 2016 . These analyses are presented in the CSR.[44]

The results from the CSR are presented within this submission and were used in the economic modelling. Where these are the same as the published data, the publications are also referenced. The results in the CSR are based on assessment by an Independent Radiology Committee (IRC) and FDA-recommended censoring rules, and the p-values presented are two-sided. The application of FDA-

recommended censoring rules for PFS necessarily reduced the number of events available for analysis. To increase the number of events that would be included in the analyses, the data cut-off for radiographic endpoints in the CSR was extended to 15 September 2016 (database extract 13 January, 2017; the latest date for which OS data were available).

13 January 2017 data cut off

Overall survival analyses as presented in the CSR were conducted with the most mature OS data available from the Alliance at the time (cut-off date of 13 January 2017).

July 2017 data cut-off

Results from the updated OS analysis published by the Alliance in September 2017 (data cut-off July 2017) are also provided.[47]

For the reasons described above, there are some discrepancies between the results presented here and those in the study publication.[2] Key differences in the analyses are summarised in Table 7. A summary of the clinical effectiveness evidence is provided in

Table 8.

Table 7 Key differences between investigator and regulatory analyses, CABOSUN

	Original report (Choueiri 2016[2])	**CSR [44] and Choueiri 2017[47]**	
Reader	**Investigator**	**Investigator**	**IRC**
No. of patients with radiographic images	157	157	156
No. of events	123	107	92
Cut-off date (PFS and ORR)	April 2016	September 2016	
Cut-off date (OS)	April 2016	January 2017 (CSR) / July 2017 (Choueiri 2017)	
Censoring rules (PFS)	Alliance	FDA guidance	
Censor for non-protocol systemic anticancer therapy	No	Yes	
Censor if event after ≥2 missing assessments	No	Yes	
Stratified analysis[a]	Yes	Yes	
P-value sided	1	2	

[a] Stratification factors were IMDC risk group (poor, intermediate) and bone metastases (yes, no).
Abbreviations: CSR, clinical study report; FDA, Food and Drug Administration; IRC, Independent

Radiology Committee; ORR, objective response rate; OS, overall survival; PFS, progression-free survival. Source: Adapted from Choueiri et al 2017[47]

Table 8 Clinical effectiveness evidence

Study	CABOSUN[2,44,47]		
Study design	Open label RCT (phase 2)		
Population	Patients with untreated clear cell metastatic RCC, ECOG performance status of 0 to 2 and intermediate or poor risk per IMDC criteria		
Intervention(s)	Cabozantinib (60 mg once per day)		
Comparator(s)	Sunitinib (50 mg once per day; 4 weeks on, 2 weeks off)		
Indicate if trial supports application for marketing authorisation	Yes	**Indicate if trial used in the economic model**	Yes
Rationale for use/non-use in the model	CABOSUN was used in the model because it is the trial on which the regulatory submission is based, and it is the only trial of cabozantinib in the population stated above. The economic model was based on the regulatory analyses as published in the CSR.		
Reported outcomes specified in the decision problem	**Progression-free survival, overall survival**, objective response rate, **adverse effects of treatment** * Outcomes used in the economic model are given in bold		
All other reported outcomes	Duration of response		

Abbreviations: CSR, clinical study report; IMDC, ,International Metastatic RCC Database Consortium; RCT, randomised controlled trial.

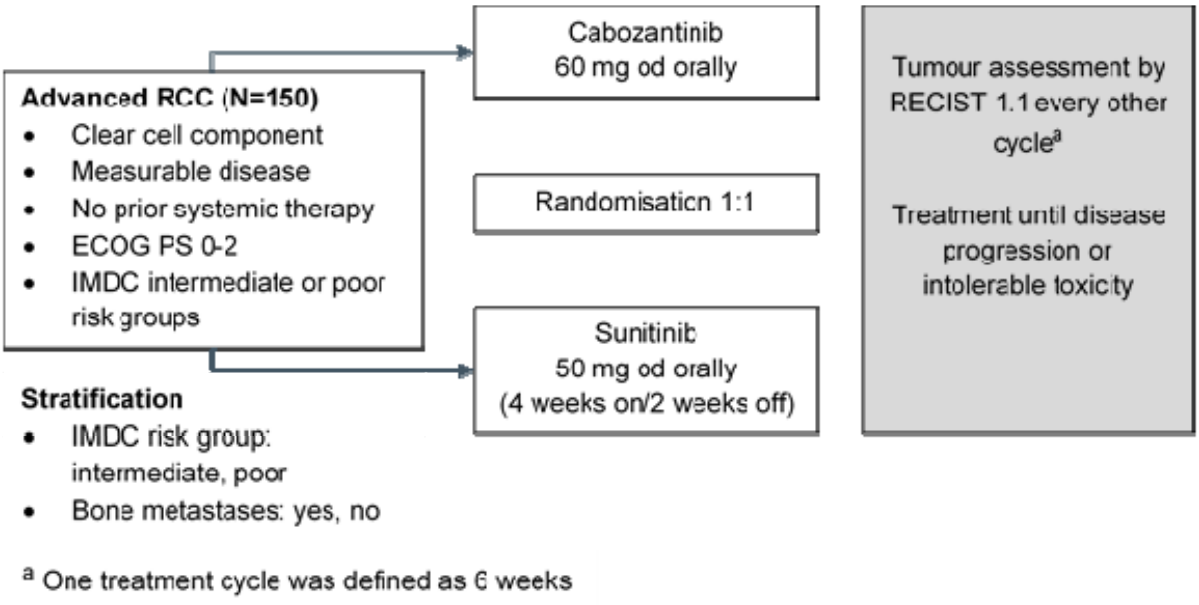
B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

Trial design

CABOSUN was an open label, phase 2, RCT comparing cabozantinib with sunitinib in patients with previously untreated advanced or metastatic RCC of poor or intermediate risk, as defined by the IMDC criteria (see Section B.1.3 for definition). The primary objective was to determine if patients treated with cabozantinib had improved PFS compared with patients treated with sunitinib. Crossover between treatment arms was not permitted. However, as would be expected, many patients received subsequent anti-cancer therapy after discontinuing study treatment (see Section B.2.6, Table 16).

Both treatments were given at their approved (or anticipated approved) dosage and schedule. The open label design was chosen in order to enable appropriate dose modifications of study treatments to manage adverse events (AEs). A blinded central review by an IRC was undertaken retrospectively to minimise bias for the PFS and response end points.[44] The study design is shown in Figure 4, and details of methodology are given in Table 9.

Figure 4 Study design



ECOG PS, Eastern Cooperative Oncology Group performance status; od, once daily; RECIST, Response Evaluation Criteria In Solid Tumors

Table 9 Summary of trial methodology

Trial number (acronym)	**CABOSUN**
Location	77 centres in the US
Trial design	Randomised, active-controlled, multicentre phase 2 open label study. **Randomisation:** randomisation was 1:1, stratified by IMDC risk group (intermediate or poor) and presence of bone metastasis (yes/no). Randomisation was carried out through the web-based Oncology Patient Enrollment Network (OPEN) registration system.
Eligibility criteria for participants	Patients were required to be ≥18 years of age and have documented RCC with some component of clear cell histology, that was advanced (defined as not amenable to curative surgery or radiation therapy) or metastatic (American Joint Committee on Cancer stage IV). Other key eligibility criteria were:

	• Intermediate or poor risk by IMDC criteria • ECOG performance status 0 to 2 • No prior systemic treatment for RCC • No active brain metastases; patients with treated brain metastases which had been stable for at least 3 months were eligible • Adequate organ and marrow function with no uncontrolled significant illness. A full list of inclusion and exclusion criteria is available in the CSR.[44]
Settings and locations where the data were collected	The study was conducted in hospital and outpatient clinics.
Trial drugs (the interventions for each group with sufficient details to allow replication, including how and when they were administered) Intervention(s) (n=[x]) and comparator(s) (n=[x]) Permitted and disallowed concomitant medication	Cabozantinib was administered orally once per day at a dose of 60 mg (n=79 randomised). Sunitinib was administered orally once per day at a dose of 50 mg for 4 weeks, followed by a 2-week break (n=78 randomised). A treatment cycle was defined as 6 weeks in both study groups. Adverse events were managed with treatment interruptions and dose reductions. Cabozantinib dose reductions were to 40 and 20 mg, and sunitinib dose reductions were to 37.5 and 25 mg. **Duration of treatment:** Treatment was continued until disease progression, intolerance to therapy, or withdrawal of consent for treatment. **Concomitant medications:** • Patients received full supportive care (including transfusions of blood and blood products, erythropoietin, antibiotics, antiemetics, and other agents), when appropriate. • Prophylactic measures were taken to prevent or reduce the severity of palmar-plantar erythrodysaesthesia syndrome (PPES; hand-foot syndrome). • Palliative radiotherapy was not permitted. • Concomitant use of medications that are strong inhibitors/inducers of CYP3A4 was to be avoided.
Primary outcomes (including scoring methods and timings of assessments)	The primary outcome was PFS, defined as the interval between randomisation and first documentation of disease progression, or death from any cause. When the study was conducted, progression was investigator-assessed. For the regulatory submission, a blinded central review of the radiographic images was carried out retrospectively by an IRC, to determine progression and response. The results reported in this submission to NICE are IRC-assessed. Progression was assessed according to RECIST 1.1 at screening and every 12 weeks (i.e. every 2 treatment cycles). Brain imaging was carried out at screening, but after randomisation it was only performed if signs or symptoms of

	brain metastases developed. A bone scan was performed at screening; after randomisation, bone scans were performed every 12 weeks but only if the baseline scan was indicative of metastases or if signs or symptoms of metastases developed.
Other outcomes used in the economic model/specified in the scope	The secondary end points were: • ORR, defined as proportion of subjects at the time of data cut-off with a best overall response of CR or PR, which was confirmed by a subsequent visit ≥ 28 days later (assessment as for PFS, above). • OS*, defined as time from randomisation to death from any cause. • AEs*, graded by 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. *OS and AEs are utilised in the model
Pre-planned subgroups	• Exploratory analyses were carried out on subgroups based on the stratification factors used in randomisation (ECOG PS, bone metastases at baseline, and IMDC risk category). • Exploratory analyses were also carried out using subgroups based on age, sex and race. These subgroup analyses were not pre-planned. See Appendix E for subgroup definitions.

Abbreviations: AE, adverse event; AJCC, American Joint Committee on Cancer; CTCAE, Common Terminology Criteria for Adverse Events; ECOG, Eastern Cooperative Oncology Group; IMDC, International Metastatic RCC Database Consortium; IRC, independent radiological committee; OS, overall survival; PFS, progression-free survival; RECIST, Response Evaluation Criteria In Solid Tumors. Source: Clinical Study Report,[44] Choueiri et al 2016[2]

Baseline patient characteristics

In total 157 patients were randomised, 79 to cabozantinib and 78 to sunitinib. Baseline demographic and clinical characteristics are summarised in Table 10. Demographic characteristics were well balanced between groups. Median age was 63 and 64 years (cabozantinib and sunitinib groups, respectively), and patients were predominantly male and of white race. Most patients (81% in each group) had intermediate risk disease and 13% in each group had poor performance status (ECOG PS 2). All patients had metastatic disease. Bone metastases were present in just over one-third of patients in each group.

Table 10 Baseline patient characteristics

Characteristic (n, %)	Cabozantinib (n=79)	Sunitinib (n=78)
Age, years		
Median (range)	63 (40-82)	64 (31-87)
Sex, male	66 (84)	57 (73)
Race		
White	70 (89)	75 (96.2)
Black	3 (4)	2 (2.6)
Asian	2 (3)	0
Other, unknown or not reported	5 (6)	1 (1)
ECOG PS		
0	36 (46)	36 (46)
1	33 (42)	32 (41)
2	10 (13)	10 (13)
IMDC risk group		
Intermediate	64 (81)	63 (81)
Poor	15 (19)	15 (19)
Bone metastases		
Yes	29 (37)	28 (36)
No	50 (63)	50 (64)
Prior nephrectomy		
Yes	57 (72)	60 (77)
No	22 (28)	18 (23)
Metastases[a]		
≥ 1 metastatic site	79 (100)	78 (100)
Visceral metastases	61 (77)	56 (72)

a, as reported by the investigator on the on-study case-report form. Abbreviations ECOG, Eastern Cooperative Oncology Group; IMDC, International Metastatic Renal Cell Carcinoma Database Consortium; PS, performance status. Source: Clinical Study Report[44], Choueiri et al. 2016[2]

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

Participant flow through the trial is shown in Appendix D.

The statistical analyses used in CABOSUN are summarised in Table 11.

Table 11 Summary of statistical analyses, CABOSUN

Hypothesis	The null hypothesis was that the hazard ratio (HR) of progression of the two treatment arms would be 1.0; the alternative hypothesis was that the HR would be 0.67, favouring the experimental arm (cabozantinib) over the control arm (sunitinib).
Analysis populations	All efficacy analyses were carried out in the intent-to-treat (ITT) population, defined as all patients who were randomised. Patients were analysed according to their assigned treatment arm. Safety analyses were carried out in the safety population, defined as all patients who received any treatment with cabozantinib or sunitinib. Patients were analysed according to actual treatment received.
Sample size and power calculation	With 123 events (progressions or deaths), the log-rank statistic had 85% power to detect an HR of 0.67 for PFS, assuming a one-sided type I error of 0.12 (equivalent to an increase in median PFS from 8 months in the sunitinib arm to 12 months in the cabozantinib arm). The one-sided test corresponded to the one-sided study hypothesis. The following assumptions were made to achieve the target of 123 PFS events: an accrual rate of 5.8 patients per month over a 24-month enrolment period, 20 months of follow-up after study closure for the PFS end point, and an exponential distribution of PFS. Allowing for a 7% ineligibility rate, the total sample size was 150 patients.
Statistical analysis of primary outcome	The primary outcome was PFS. The Kaplan-Meier method was used to estimate the median PFS and associated 95% confidence intervals (CIs). The stratified and unstratified HRs and corresponding 95% CIs were estimated using a Cox regression model. A stratified log-rank statistic was used to compare the two treatment arms, and two-sided p-values were presented in the CSR for regulatory purposes. Analyses of PFS were performed based on FDA-recommended rules. (In this submission we present the FDA-rule-based censoring and two-sided values from the CSR, as used in the regulatory submission.)
Statistical analysis of other outcomes	No adjustments were made for multiplicity in the analysis of secondary endpoints. **Overall survival (OS)** The statistical methods used for the OS analyses were the same as those described for the analyses of the primary PFS endpoint (i.e., Kaplan-Meier, log-rank test, Cox regression). For subjects who were alive at the time of data cut-off or were permanently lost to follow-up, duration of OS was censored at the earliest of date of withdrawal of consent from all follow-up, data cut-off date, or the date the subject was last known to be alive. **Objective response rate (ORR)** Point estimates of ORR with CIs were calculated using exact methods. The difference in response rates between the treatment arms and associated CIs was also calculated; 2-sided p-values were generated using the Fisher exact test to compare the two treatment arms. Analyses using the Cochran-Mantel-Haenszel method to adjust for

	the randomisation stratification factors were also presented. The odds ratio and corresponding p-value was also calculated from Proc logistic. Waterfall plots showed the best percentage change in target lesion size as assessed by the IRC or the Investigator.
Interim analyses	The data were analysed at several different time points, as shown in Section B.2.2, Table 7.
Treatment of missing data	In the ORR analysis, subjects who did not have any post baseline tumour assessments were counted as non-responders. Censoring rules for the time-to-event analyses were applied in accordance with FDA guidance.[48] In the retrospective independent radiology committee assessment of PFS and ORR, no values were imputed for patients for whom a complete set of baseline and post-baseline radiographic images were not available.

Abbreviations: FDA, Food and Drug Administration; HR, hazard ratio; ITT intent-to-treat; PFS, progression-free survival; CSR clinical study report; OS, overall survival; ORR, objective response rate; IRC, independent radiology committee. Source: Clinical Study Report[44], Choueiri et al. 2016[2]

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

The complete quality assessment for each trial is provided in Appendix D.

B.2.6 Clinical effectiveness results of the relevant trials

Summary

- Evidence on the efficacy and safety of cabozantinib in previously untreated advanced RCC is available from one RCT, the CABOSUN trial, which compared cabozantinib with sunitinib in adults with previously untreated RCC patients with intermediate or poor risk per IMDC criteria.

Primary endpoint

- Cabozantinib treatment resulted in a clinically meaningful and statistically significant prolongation of PFS compared with sunitinib, by IRC assessment (8.6 vs. 5.3 months, difference 3.3 months, HR 0.48 [95% CI 0.31, 0.74], p=0.0008).[44,47]
- The PFS benefit obtained with cabozantinib was not diminished in patients who had bone metastases: the HRs in patients with and without bone metastases were 0.51 (95% CI 0.26, 0.99) and 0.50 (95% CI 0.29, 0.85), respectively.[44,47]

Secondary endpoint

- Cabozantinib increased median OS by 9.3 months compared with sunitinib, although the data are immature and statistical significance was not reached (HR adjusted for stratification factors 0.74 [95% CI 0.47, 1.14]; stratified 2-sided log-rank p-value = 0.1700).[44]
- ORR with cabozantinib was twice that observed with sunitinib by IRC assessment (20% [95% CI 12.0, 30.8] vs. 9% [95% CI 3.7, 17.6]).[44,47]

The data presented here are taken from the CSR[44] and reflect those used in the regulatory submission. The trial results were initially published by the Alliance (Choueiri et al. 2016[2]), PFS and ORR data based on IRC assessment and FDA censoring rules along with an updated OS analysis were also presented by Choueiri et al. 2017.[47]

The data from the CSR are used in the economic modelling.

Primary end point

Progression-free survival by IRC assessment

Median follow-up for PFS was 25.0 months. A complete set of baseline and post-baseline radiographic images was available for 143 patients. The remaining 13 patients were not assessed due to the following reasons: baseline images available but no post-baseline tumour assessments performed for 11 patients (1 cabozantinib, 10 sunitinib), incomplete set of baseline images collected for 1 patient (cabozantinib), and corrupt baseline images for 1 patient (sunitinib). See Table 12 for details of censoring.

- Cabozantinib was associated with a statistically significant and clinically meaningful improvement in PFS compared with sunitinib (8.6 vs. 5.3 months, difference 3.3 months, HR 0.48 [95% CI 0.31, 0.74], p=0.0008).[44,47] PFS results by IRC are shown in Table 12 and Figure 5.
- PFS by investigator assessment, using FDA-recommended censoring rules, was similar to the IRC analysis. The Kaplan-Meier estimates for median PFS were 8.3 vs. 5.4 months for the cabozantinib and sunitinib arms, respectively (HR adjusted

for stratification factors =0.56 [95% CI 0.37, 0.83]; stratified 2-sided log-rank p-value = 0.0042).[44,47]

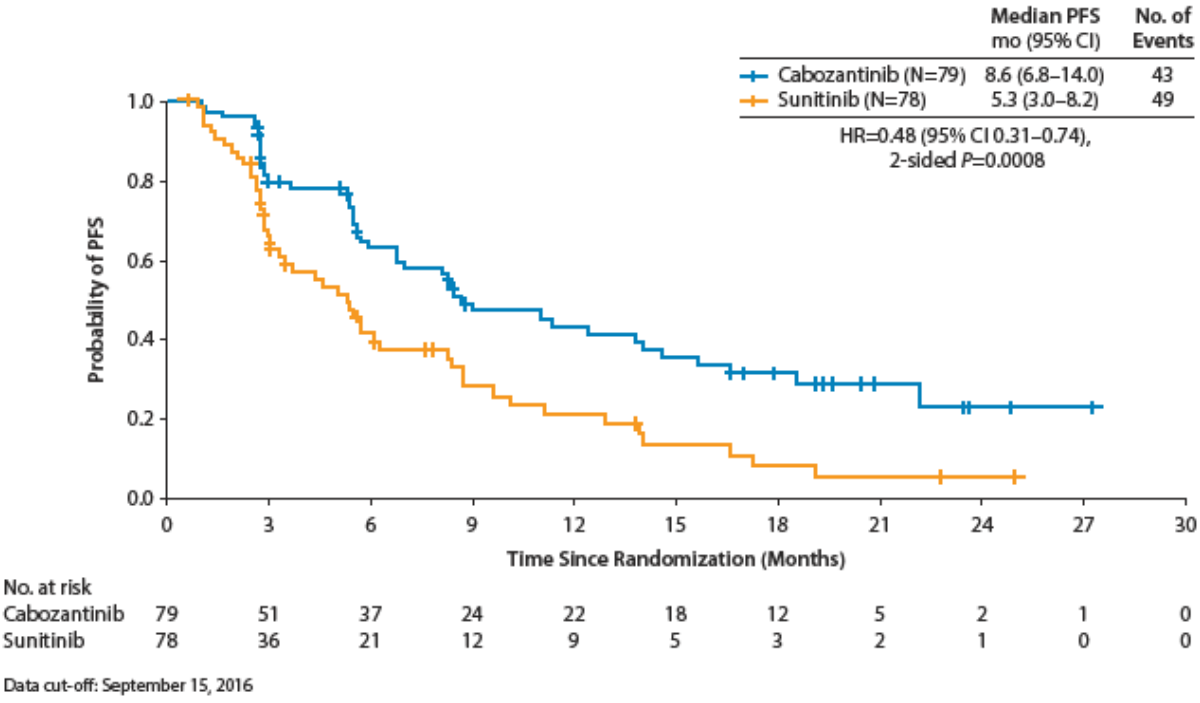
- The PFS benefit obtained with cabozantinib was not diminished in patients who had bone metastases: the HRs in patients with and without bone metastases were 0.51 (95% CI 0.26, 0.99) and 0.50 (95% CI 0.29, 0.85), respectively[44,47] (see Appendix E for full results of the subgroup analyses).

Table 12 Progression-free survival by IRC, ITT population

N (%)	Cabozantinib (N = 79)	Sunitinib (N = 78)
Censored	36 (46)	29 (37)
2 or more missed ATAs prior to event	5 (6)	0
No baseline and post-baseline ATAs	0	2 (3)
No event by last ATA	10 (13)	3 (4)
No post-baseline ATA	1 (1)	6 (8)
Systemic anticancer therapy	20 (25)	18 (23)
Event		
Death	3 (4)	6 (8)
Documented progression	40 (51)	43 (55)
Median PFS (95% CI), months	8.6 (6.8, 14.0)	5.3 (3.0, 8.2)
Two-sided log-rank p-value: stratified [b]	0.0008	
Hazard ratio (95% CI); stratified [b,c]	0.48 (0.31, 0.74)	
PFS at 12 months (% event-free)	43.1	21.1

a, Median and percentiles are based on Kaplan-Meier survival estimates; b Stratification factors were IMDC risk categories (intermediate risk, poor risk) and bone metastasis (yes, no); c, Estimated using the Cox proportional hazard model adjusted. Abbreviations: ATA, adequate tumour assessment; CI, confidence interval; IRC, independent radiology committee; PFS, progression-free survival. Hazard ratio < 1 indicates PFS in favour of cabozantinib. Source: Clinical Study Report[44], Choueiri et al 2017[47]

Figure 5 Kaplan-Meier plot of PFS by IRC, ITT population



Abbreviations: CI, confidence interval; HR, hazard ratio; PFS, progression-free survival. Source: Clinical Study Report[44], Choueiri et al. 2017[47]

Secondary end points

Overall survival (data cut-off 13 January 2017)

At the data cut-off point used for analysis of OS in the CSR (13 January 2017), median time of follow-up for OS was 28.9 months. The survival data were immature at this analysis – 38 deaths were recorded in the cabozantinib arm and 45 in the sunitinib arm.[44]

- Cabozantinib was associated with a substantial numerical improvement in median OS compared with sunitinib. The Kaplan-Meier estimates for median OS were 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. However, the immaturity of the data meant that there was a notable degree of censoring around the median estimates, and they should therefore be interpreted with caution. The HR adjusted for stratification factors was 0.74 (95% CI 0.47, 1.14; stratified 2-sided log-rank p-value = 0.1700).[44]
- Results of landmark analysis demonstrate that at 30 months (the longest landmark analysis reported) the percentage of patients event-free were 50.7% and 30.3% for cabozantinib and sunitinib, respectively.[44]

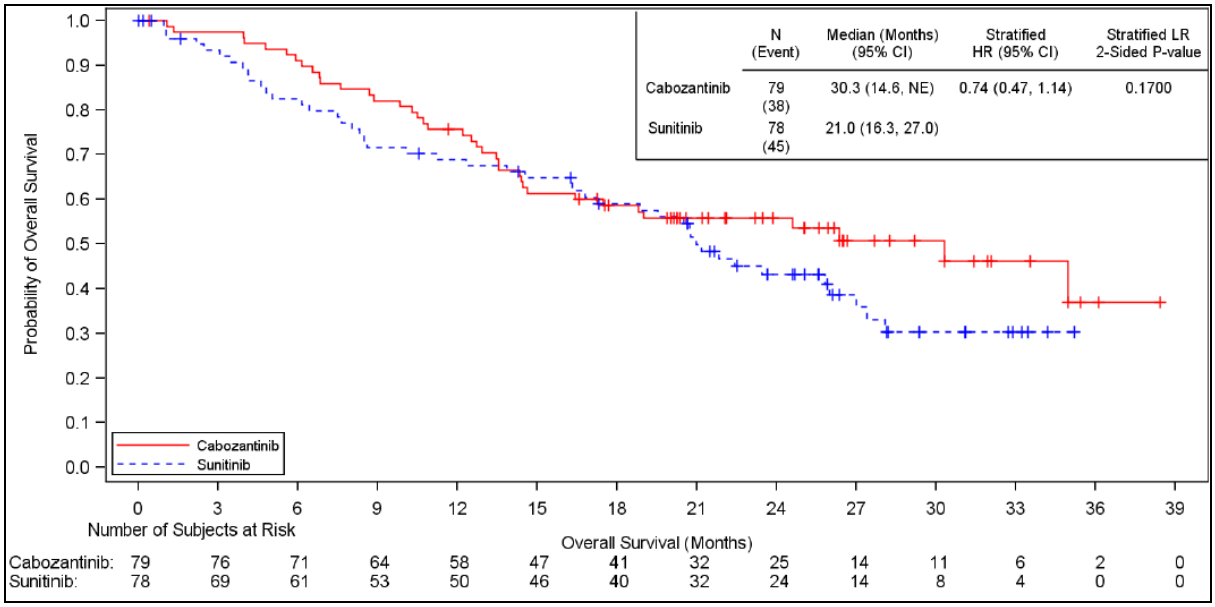
OS data are summarised in Table 13 and Figure 6. (Note: these data are used to inform the economic model.)

Table 13 Overall survival, ITT population, data cut-off 13 January 2017

	Cabozantinib (N = 79)	Sunitinib (N = 78)
Number (%) of patients		
Censored (no event as of cut-off date)	41 (52)	33 (42)
Deaths	38 (48)	45 (58)
Duration of OS, months (median [95% CI])[a]	30.3 (14.6, NE)	21.0 (16.3, 27.0)
2-sided stratified log-rank p-value [b]	0.1700	
Stratified hazard ratio (95% CI)[b, c]	0.74 (0.47, 1.14)	
% event-free at 30 months	50.7	30.3

a, median based on Kaplan-Meier survival estimate; b, stratification factors were IMDC risk category (intermediate, poor) and bone metastases (yes, no); c, Estimated using Cox proportional hazards model; HR <1 indicates OS in favour of sunitinib. Abbreviations: CI, confidence interval; ITT, intention to treat; NE, not estimable; OS, overall survival. Source: Clinical Study Report[44]

Figure 6 Kaplan-Meier plot of overall survival, ITT population, data cut-off 13 January 2017



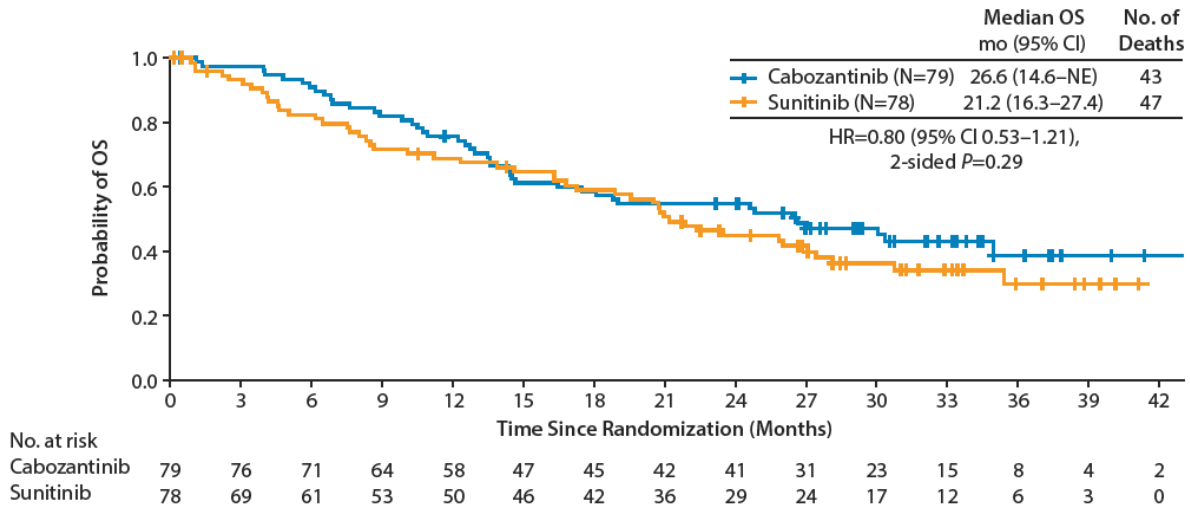
Stratification factors were IMDC risk category (intermediate, poor) and bone metastases (yes, no). Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; LR, log rank; NE, not estimable. Source: Clinical Study Report[44]

Updated overall survival results (data cut-off 1 July 2017)

Updated OS results with a data cut-off of 1 July 2017 were published by the Alliance (Figure 7).[47] Median OS continued to favour cabozantinib, at 26.6 (95% CI 14.6, NE) vs. 21.2 (95% CI 16.3, 27) months for cabozantinib and sunitinib, respectively. However, the confidence intervals around the medians and HR remained wide due

to the relatively low number of deaths, and the HR (0.80 in favour of cabozantinib, 95% CI 0.53, 1.21, 2-sided p-value = 0.29) did not reach statistical significance.

Figure 7 Overall survival, data cut-off 1 July 2017



Abbreviations: CI, confidence interval; HR, hazard ratio; NE, not estimable; OS, overall survival.
Source: Choueiri et al 2017[47]

Objective response rate by IRC

The ORR with cabozantinib was twice that seen with sunitinib. Cabozantinib also resulted in a markedly higher disease control rate. Radiographic images were collected for 156 of 157 patients. Response is summarised in Table 14, and waterfall plots are shown in Figure 8.

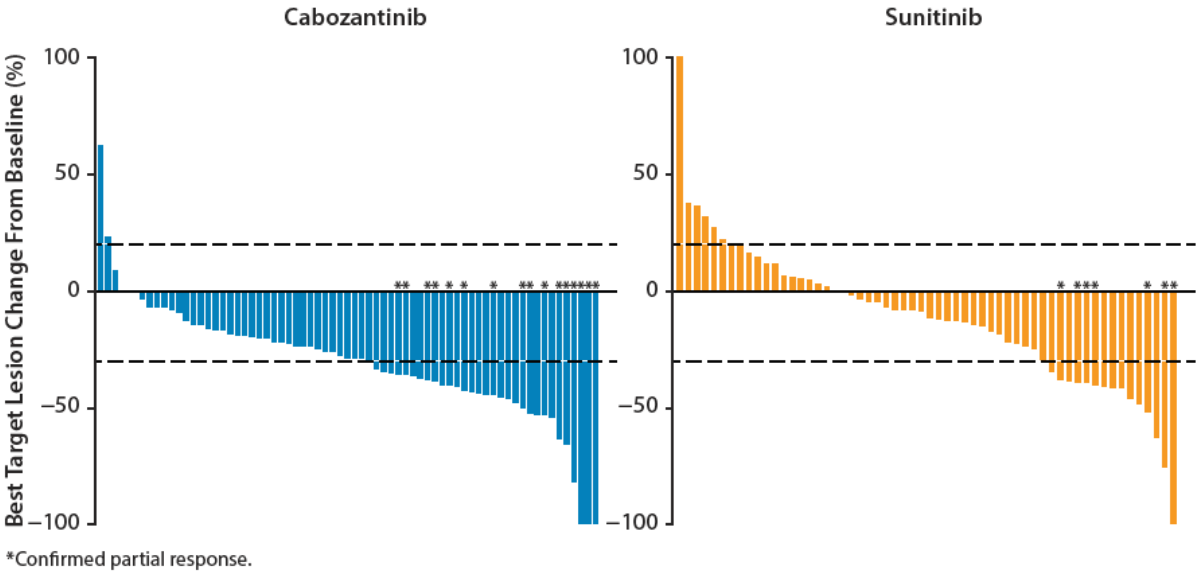
- The IRC-determined ORR was 20% (95% CI 12.0, 30.8) in the cabozantinib arm, compared with 9% (95% CI 3.7, 17.6) in the sunitinib arm (ITT population). All responses were partial responses. The difference in ORR between treatment arms was 11.3% (95% CI 0.4, 22.2, stratified 2-sided p-value = 0.0406).[44,47]
- The disease control rate (complete response + partial response + stable disease) was markedly higher with cabozantinib than with sunitinib, at 75% and 47%, respectively.[47]

Table 14 Tumour response by IRC, ITT population

N (%)	Cabozantinib (N = 79)	Sunitinib (N = 78)
Best overall response		
Confirmed complete response	0	0
Confirmed partial response	16 (20)	7 (9)
Stable disease	43 (54)	30 (38)
Progressive disease	14 (18)	23 (29)
Unable to evaluate	4 (5)	6 (8)
Missing[a]	2 (3)	12 (15)
Objective response rate		
N (%)	16 (20)	7 (9)
95% CI, %	(12.0, 30.8)	(3.7, 17.6)
Treatment difference	11.3 (0.4, 22.2)	
Stratified CMH test 2-sided p-value[b]	0.0406	

a, Patients who were unable to evaluate/missing had the following end of treatment reasons: Cabozantinib - Discontinuation due to AE (5), withdrew consent prior to treatment (1); Sunitinib - Discontinuation due to AE (6), death on study (2), disease progression during treatment (1), withdrew consent prior to treatment (5), withdrew consent after starting treatment (4). b, Stratification factors were IMDC risk group and bone metastases. Abbreviations: CI, confidence interval; CMH, Cochran–Mantel–Haenszel. Source: Clinical Study Report[44]

Figure 8 Waterfall plot of best percentage change in tumour size, by IRC, ITT population



Six patients in the cabozantinib group and 18 patients in the sunitinib group were not evaluable because they had no adequate post-baseline imaging assessments. Abbreviations: IRC, independent radiology committee; ITT, intention to treat. Source: Choueiri et al. 2017[47]

Summary of PFS and response results

Table 15 below, published by Choueiri et al.,[47] shows a summary of the analyses performed for the radiographic end points (PFS and response) by investigator and IRC assessment.

Table 15 Summary of PFS and tumour response analyses for CABOSUN (results in bold were used in the economic model)

	Original report[a2] Investigator-assessed, Alliance censoring rules		**September 2016 cut-off (reported in CSR) Investigator-assessed, FDA censoring rules**		**September 2016 cut-off (reported in CSR) IRC-assessed, FDA censoring rules**	
	Cabozantinib (N=79)	**Sunitinib (N=78)**	**Cabozantinib (N=79)**	**Sunitinib (N=78)**	**Cabozantinib (N=79)**	**Sunitinib (N=78)**
Progression-free survival						
Median PFS, months	8.2	5.6	8.3	5.4	**8.6**	**5.3**
Stratified HR (95% CI)	0.66 (0.46-0.95)		0.56 (0.37-0.83)		**0.48 (0.31-0.74)**	
P value	0.012 (1-sided)		0.0042 (2-sided)		**0.0008 (2-sided)**	
Tumour response						
Objective response rate (95% CI),[b] %	33 (23-44)	12 (5-21)	33 (23-44)	12 (5-21)	20 (12-31)	9 (4-18)
Disease control rate,[c] %	78	54	76	49	75	47
Progressive disease,[d] %	18	26	18	24	18	29
Not evaluable or missing, %	4	21	6	27	8	23
Any reduction in target lesion, %	87	44	85	38	80	50

a, Data cut-off: April 11, 2016; b, One complete response was observed with cabozantinib for both investigator assessments, and one complete response was observed with sunitinib for the original investigator assessment; all other responses were partial responses; c, Complete response + partial response + stable disease; d, Progressive disease as best overall response. Abbreviations: CI, confidence interval; FDA, Food and Drug Administration; HR, hazard ratio. Source: Choueiri et al. 2017[47]

Subsequent anti-cancer treatment

The protocol did not allow for crossover from one study treatment to the other within the study, and non-protocol anticancer therapy was not to be initiated until after study treatment had been discontinued.[44] Subsequent systemic non-radiation

anticancer therapy was used by a similar proportion of patients in each arm: 57% and 58% in the cabozantinib and sunitinib arms, respectively (ITT population). The median time to first systemic non-radiation anti-cancer therapy was longer in the cabozantinib arm: median times were 196 (range 56, 877) days in the cabozantinib arm and 147 (4, 725) days in the sunitinib arm.[44] When all reported subsequent therapies are included, just under half of all patients received a VEGFR-targeted TKI as a subsequent treatment. A further 13% and 15% received an anti-PD-1 or anti-PDL-1 agent (cabozantinib and sunitinib arms, respectively). Temsirolimus and everolimus were also used. Radiation therapy was used by 13% and 18%, respectively.

Details of the first subsequent anti-cancer therapy, as given in the original study publication, are shown in Table 16. Details of all subsequent anti-cancer treatments, as reported in the CSR, are given in Appendix L.

Table 16 First subsequent anti-cancer therapy, as reported by the Investigators

	No. (%)	
Therapy	**Cabozantinib (n = 79)**	**Sunitinib (n = 78)**
Any subsequent anticancer therapy	41 (51.9)	47 (60.3)
Radiotherapy	6 (7.6)	14 (17.9)
Surgery	5 (6.3)	3 (3.8)
Systemic subsequent anticancer therapy*	37 (46.8)	45 (57.7)
Axitinib	13 (16.5)	13 (16.7)
Pazopanib	10 (12.7)	4 (5.1)
Sunitinib	7 (8.9)	10 (12.8)
Temsirolimus	5 (6.3)	2 (2.6)
PD-1 inhibitors	5 (6.3)	5 (6.4)
Nivolumab	3 (3.8)	2 (2.6)
Everolimus	4 (5.1)	13 (16.7)
Sorafenib	1 (1.3)	2 (2.6)
Bevacizumab	0 (0)	5 (6.4)
Cabozantinib	0 (0)	3 (3.8)

Data are as of September 15, 2016. PD-1, programmed death 1. *First anticancer therapy reported after cessation of protocol therapy is provided. Source: Choueiri et al 2016[2]

B.2.7 Subgroup analysis

Subgroup analyses of PFS were carried out for the stratification variables (IMDC risk category and presence/absence of bone metastases) MET status (positive/negative/missing) and a range of other demographic and clinical characteristics. Results are summarised in Appendix E.

- For IRC-determined PFS there was a consistently favourable effect for cabozantinib compared with sunitinib for all the larger subgroups (≥ 20 patients per treatment arm).[44]
- The PFS benefit obtained with cabozantinib was not diminished in patients who had bone metastases: the HRs in patients with and without bone metastases were 0.51 (95% CI 0.26, 0.99) and 0.50 (95% CI 0.29, 0.85), respectively.[44,47]
- For the smaller subgroups, the numbers of patients were too low to allow for any meaningful interpretation of the results.

B.2.8 Meta-analysis

Not applicable (only one trial giving evidence on cabozantinib in untreated RCC was identified).

B.2.9 Indirect and mixed treatment comparisons

Full details of the methodology for the indirect treatment comparison are included in Appendix D.

Identification and selection of studies

A systematic literature review was carried out as described in Section B.2.1. Seven publications referring to two studies were included in the systematic literature review and considered for potential inclusion in the network meta-analysis (NMA) or indirect treatment comparison (ITC). Multiple publications reporting the same study were identified and grouped as associated references (Table 17). For the CABOSUN study the CSR[44] and patient-level data were additionally available, corresponding to final data cut off 13 January 2017 for OS and 15 September 2016 for PFS.[44] The full reference list of included and excluded records is provided in Appendix D.

Table 17 Trial overview and study mapping

Study	Full reference	Main publication (main/abstract/ additional)	Comment
ALLIANCE/A031203/CABOSUN trial; NCT01835158 **Study arms: cabozantinib; sunitinib**			
Choueiri_2016[2]	Choueiri, T. K., Halabi, S., Sanford, B. L., et al. Cabozantinib versus sunitinib as initial targeted therapy for patients with metastatic renal cell carcinoma of poor or intermediate risk: The alliance A031203 CABOSUN trial. Journal of Clinical Oncology. 2017. 35(6): p. 591-597.	Main publication Final analysis	
Choueiri_2016b[45]	Choueiri, T. K., Halabi, S., Sanford, B., et al. PR CABOzantinib versus SUNitinib (CABOSUN) as initial targeted therapy for patients with metastatic renal cell carcinoma (mRCC) of poor and intermediate risk groups: Results from ALLIANCE A031203 trial. Annals of Oncology. Conference: 41st European Society for Medical Oncology Congress, ESMO. 2016. 27(no pagination)	Abstract Refers to Choueiri_2016[2]	
Chopra_2016[46]	Chopra, M. Annual Congress of the European Society for Medical Oncology (ESMO): Copenhagen, Denmark; 7-11 October 2016. Targeted Oncology. 2016. 11(6): p. 705-709.	Abstract Refers to Choueiri_2016[2]	
COMPARZ; NCT00720941 **Treatment arms: pazopanib, sunitinib**			
Motzer_2013[27]	Motzer, R. J., Hutson, T. E., Cella, D., et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. New England Journal of Medicine. 2013. 369(8): p. 722-731.	Main publication	Overall results cut of May 2012
Guo_2013[49]	Guo, J., Jin, J., Huang, Y., et al. Comparison of PFS and safety for Asian compared to North American and European populations in the phase III trial of pazopanib versus sunitinib in patients with treatment-naive RCC (COMPARZ). Journal of Clinical Oncology. 2013. 31(6 suppl. 1).	Abstract	Abstract to Motzer_2013, no additional results of interest - not extracted, Includes ethnicity-based subgroup analysis of

			efficacy and safety
Motzer_2012[50]	Motzer, R., Hutson, T. E., Reeves, J., et al. Randomized, open-label, phase III trial of pazopanib versus sunitinib in first-line treatment of patients with Metastatic Renal Cell Carcinoma (MRCC): Results of the COMPARZ trial. Annals of Oncology. 2012. 23(p. ix13.	Abstract	Abstract to Motzer_2013, no additional results of interest - not extracted
Motzer_2014[51]	Motzer, R. J., Hutson, T. E., McCann, L., et al. Overall survival in renal-cell carcinoma with pazopanib versus sunitinib. New England Journal of Medicine. 2014. 370(18): p. 1769-1770.	Main	Letter, final OS Sept 2013
ct.gov[52]	https://ClinicalTrials.gov/show/NCT00720941	ct.gov	

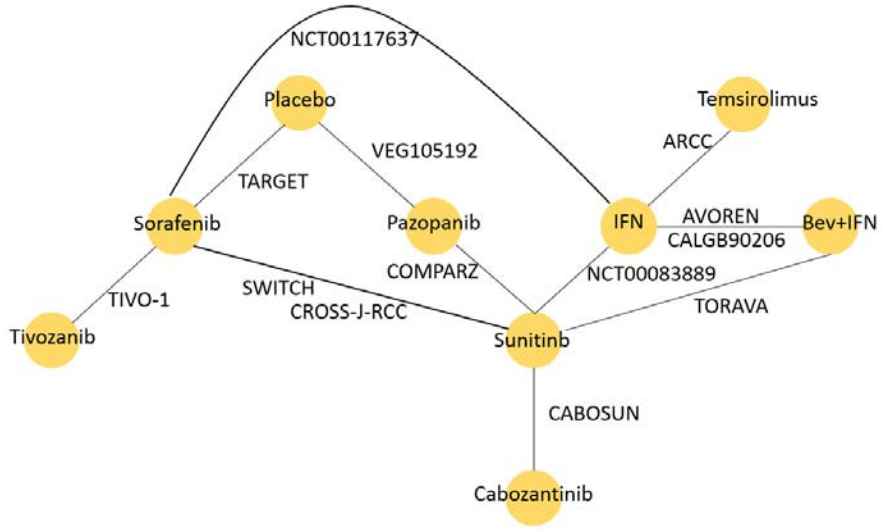
Summary of trials

In total 19 studies were identified for potential inclusion in the ITC, including studies on medicines outside of scope of this appraisal (Table 18). The potential network developed from these studies is shown in Figure 9. The full set of identified studies (including those that were outside the scope of this NICE appraisal) were checked for additional connections between cabozantinib and pazopanib.

Table 18 Primary RCT data sources included in the network evidence base

Study name	Main publication	Treatment arms
Alliance A031203 CABOSUN; NCT01835158	Choueiri_2016[2]	cabozantinib sunitinib
NCT00098657; NCT00083889	Motzer_2007b[28], Motzer_2009[53]	sunitinib interferon alpha
COMPARZ; NCT00720941	Motzer_2013[27] Motzer_2014[51]	pazopanib sunitinib
CROSS-J-RCC; NCT01481870	Tomita_2014[54] Tomita_2017[43]	sunitinib sorafenib
NCT00334282; VEG105192	Sternberg_2013[55] Sternberg_2010[56]	pazopanib placebo
NCT00117637	Escudier_2009[57]	sorafenib interferon alfa-2a
TARGET	Negrier_2010[58]	sorafenib placebo
AVOREN	Escudier_2007c[59] Escudier_2010[60]	bevacizumab + interferon alfa placebo + interferon alfa
CALGB 90206	Rini_2008a[61] Rini_2010[62]	bevacizumab + interferon alfa placebo + interferon alfa-2a
TIVO-1; NCT01030783	Motzer_2013[63]	tivozanib sorafenib
TORAVA; NCT00619268	Negrier_2011[64]	temsirolimus+bevacizumab* interferon alfa plus bevacizumab sunitinib
Global ARCC; NCT00065468	Hudes_2007[65]	temsirolimus interferon alfa interferon alfa + temsirolimus*
SWITCH	Eichelberg_2015[66]	sunitinib; sorafenib sorafenib; sunitinib
*Treatment arms not of interest in both global SLR and NICE scope		

Figure 9 Primary evidence network for potential network meta-analysis



Since the ITC for this appraisal of cabozantinib only needed to include the comparators relevant to the scope (i.e. pazopanib and sunitinib), studies which did not include these comparators were therefore excluded, unless they provided an intermediate link. The studies included in the final evidence base utilised for the ITC are summarised in Table 19. Quality assessments of each study are provided in Appendix D.

Table 19 Studies included in the final evidence base for indirect treatment comparison.

Study	Design	Population	Treatment arm	Primary endpoint
Alliance A031203 CABOSUN; NCT01835158	Phase 2 RCT Open label Multi centre	Adult patients with advanced or metastatic RCC with no prior systemic therapy	Cabozantinib Sunitinib	PFS
COMPARZ; NCT00720941	Phase 3b RCT Open label Multi centre	Adult patients with advanced or metastatic RCC with no prior systemic therapy	Pazopanib Sunitinib	PFS
Abbreviations: PFS, progression-free survival; RCC, renal cell carcinoma; RCT, randomised controlled trial				

Feasibility assessment

The ITC was planned on two efficacy endpoints: OS and PFS. These represent key outcomes of interest to clinicians and patients and are consistently selected as primary and secondary efficacy endpoints in RCC trials. The outputs of the ITC for these efficacy endpoints are utilised in the health economic analysis presented in Section B 3 (cost-effectiveness).

Feasibility assessment was carried out to assess whether there were differences in study and patient characteristics across comparisons that could affect the treatment effects. In addition, differences within or between direct treatment comparisons were also checked. The heterogeneity assessment aimed to determine to which extent the study results could be combined into an ITC and it was based on the inclusion/exclusion criteria used in the literature review. A publication was excluded from the analysis if it did not provide PFS or OS as an efficacy outcome either in the format of HR or Kaplan-Meier data. The studies were compared with regard to trial inclusion/exclusion criteria and baseline patient characteristics such as prognostic factors and risk score, previous therapies (if any) and disease stage. Specifically, we investigated whether the studies identified in the systematic review classified patients according to their prognostic risk level at baseline: favourable, intermediate or poor. The classification was based either on the MSKCC model or the IMDC

model. The data from studies meeting these more stringent ITC inclusion criteria were analysed with respect to the therapeutic efficacy of different drugs for prolonging PFS or OS, allowing construction of the PFS and OS-specific networks.

The focus of the analysis was the comparison of treatment effects in terms of OS HR and PFS HR in ITT patients, intermediate prognostic risk patients and poor prognostic risk patients. In the base case stratified and/or cross-over adjusted OS HRs were used when available. PFS could be measured by an IRC or by the investigators (INV). The IRC assessment of disease progression was deemed likely to lead to the least biased estimates, and hence it was prioritised, if available. The INV-assessed PFS was considered only where IRC-assessed PFS was not available.

Data availability assessment

In order to assess the feasibility of performing an ITC, data availability for OS and PFS HRs and Kaplan-Meier curves were first assessed (see Table 20 and Table 21). The focus of the feasibility assessment was on the results for ITT patients, as well as subgroups of intermediate and poor prognostic risk patients, since the CABOSUN trial excluded patients with favourable prognostic risk. For ITT patients, availability of adjusted/stratified and unadjusted/un-stratified OS HRs were assessed. The availability of PFS HRs evaluated by INV or IRC was also separately assessed.

Table 20 Availability of HR and KM plots for OS

Trial Name/ Author, year	Arm	OS ITT adjusted/stratified		OS ITT unadjusted/ un-stratified		OS Intermediate risk category		OS Poor risk category	
		HR (95% CI)	KM	HR (95% CI)	KM	HR (95% CI)	KM	HR (95% CI)	KM
CABOSUN CSR[44]	Cabo Suni	0.74 (0.47, 1.14)	Fig. 4	NA	NA	0.80 (0.49, 1.31)[44]	NA	0.51 (0.20, 1.32)[44]	NA
COMPARZ Motzer et al. 2014[51]	Pazo Suni	0.92 (0.79,1.06)	Fig. 1	NA	NA	0.9 (0.74, 1.09)	NA	0.85 (0.56, 1.28)	NA

Abbreviations: Cabo, cabozantinib; Suni, sunitinib; Pazo, pazopanib; ITT, intent-to-treat; OS, overall survival; PFS, progression-free survival; HR, hazard ratios; KM, Kaplan-Meier curves: NA, not available

Table 21 Availability of HR and KM plots for PFS

Trial Name/ Author, year	Arm	PFS ITT Independent review committee		PFS ITT Investigator assessed		PFS Intermediate risk category		PFS Poor risk category	
		HR (95% CI)	KM	HR (95% CI)	KM	HR (95% CI)	KM	HR (95% CI)	KM
CABOSUN CSR[44]	Cabo Suni	0.48 (0.31, 0.74)	Fig. 2	0.56 (0.37,0.83)	Fig. 3	0.52 (0.33, 0.82)[44]	NA	0.31 (0.11, 0.92)[44]	NA
COMPARZ Motzer et al. 2013[27]	Pazo Suni	1.05 (0.90, 1.22)	Fig. 1	1.00 (0.86, 1.15)	NA	NA	NA	NA	NA
Abbreviations: Cabo, cabozantinib; Suni, sunitinib; Pazo, pazopanib; ITT, intent-to-treat; OS, overall survival; PFS, progression-free survival; HR, hazard ratios; KM, Kaplan-Meier curves; NA, not available.									

Assessment of heterogeneity

Qualitative assessment

A key consideration for any ITC is whether the studies identified are suitably homogeneous to facilitate a reliable comparison. This similarity comparison is achieved by comparing selected data from candidate studies. The similarity of the studies in each network was assessed based on study design and patient baseline characteristics (Table 22).

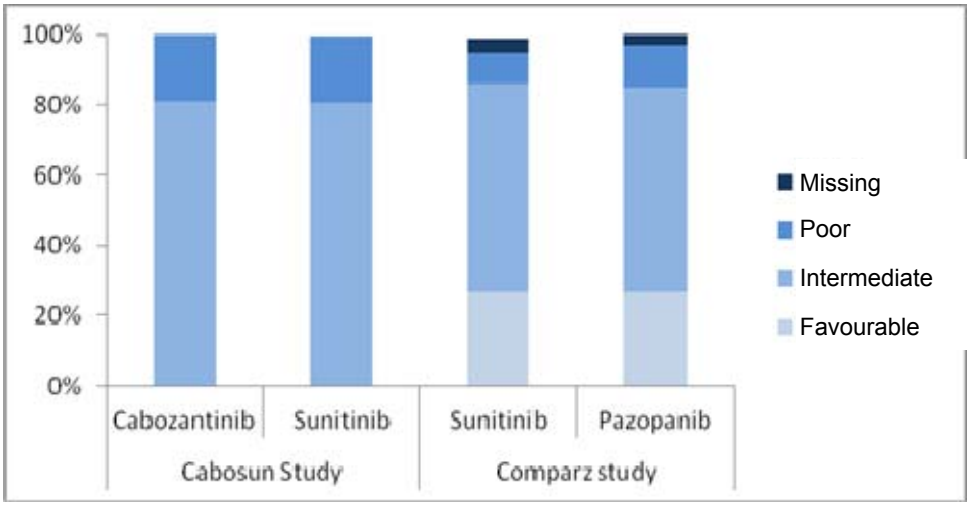
The characteristics of patients at baseline in the two identified studies were first investigated and compared. Assessment of the ECOG performance status and the MSKCC (or IMDC) status is summarised in Table 22. Compositions of risk category for each trial are presented in Figure 10. Patients had similar median age, which ranged from 61 to 64 years old. Most of the patients included were male. For the CABOSUN trial, most of the patients (92%) were white. In the COMPARZ trial, 64% of the patients were white and 34% Asian. ECOG performance status (PS) was available only for the CABOSUN trial, in which most of the patients had PS of 0 or 1 (46% had PS 0, 41% had PS 1 and 13% had PS 2). For the COMPARZ trial, PS was reported using the Karnofsky index and was 70 or 80 in 24% and 90 or 100 in 76% of patients. Regarding risk categorisation, COMPARZ reported MSKCC category and CABOSUN reported IMDC category. More details on study design and a full assessment of patient baseline characteristics are presented in Appendix D.

Table 22 Assessment of similarity between identified studies and availability of outcomes and subgroup results

Trial	Study type	Arm	MSKCC risk category N (%)				ECOG performance status N (%)			
			Favourable	Intermediate	Poor	Missing	0	1	2	Missing
CABOSUN[2] *	Phase 2 RCT Open label	Cabozantinib	0 (0)	64 (81.0)	15 (19.0)	0(0)	36 (45.6)	33 (41.8)	10 (12.7)	0 (0)
		Sunitinib	0 (0)	63 (80.8)	15 (19.2)	0 (0)	36 (46.2)	32 (41.0)	10 (12.8)	0 (0)
Comparz[27,51]	Phase 3 RCT Open label	Sunitinib	152 (27)	328 (59)	52 (9)	21 (4)	Not provided			
		Pazopanib	151 (27)	322 (58)	67 (12)	17 (3)				

Abbreviations: ECOG, Eastern Cooperative Oncology Group; MSKCC, Memorial Sloane-Kettering Cancer Centre; RCT, randomised controlled trial;
* CABOSUN reported IMDC (International Metastatic Renal Cell Carcinoma Database Consortium) risk categorisation not MSKCC

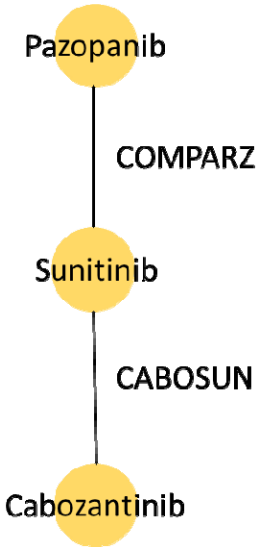
Figure 10 MSKCC/IMDC risk category composition for patients at baseline



The final network utilised in the ITC, based on the review of available data, is presented in

Figure 11.

Figure 11 Evidence network for OS and PFS



Quantitative assessment

Due to the small size of the network and its geometry with one study per comparison, a quantitative heterogeneity assessment such as the ones using the local and global Higgins coefficient was not possible.[67] We therefore performed a

qualitative heterogeneity assessment based on a side-by-side comparison of patient inclusion and exclusion criteria, baseline characteristics, and of the risk of bias of all studies included in the network.

ITC methodology

In the ITC two potential methods were considered for comparing OS and PFS endpoints: one based on the HRs and the other on the parametric curves (Kaplan-Meier). Because HR proportionality (see Appendix D) was violated for OS and PFS, an alternative method based on parametric curves was used to compare OS and PFS under different treatments. The details of ITC methods are provided in Appendix D.

ITC results

We compared cabozantinib, sunitinib and pazopanib using several models as explained in the ITC methods section in Appendix D.

Ouwens model

Survival curves (Weibull, Gompertz, log-logistic, log-normal and exponential) based on the fixed effect (FE) and random effect (RE) model are presented for OS and PFS in Appendix D (Indirect treatment comparison results, supplementary data). Model fit statistics for both OS and PFS are presented in Table 23 and Table 24.

Table 23 Fit statistics - Ouwens models

Endpoints	**OS**		**PFS**	
Model	**Fixed effects ITC**	**Random effects ITC**	**Fixed effects ITC**	**Random effects ITC**
Weibull	1757.6	1757.2	1945.8	1945.5
Log-logistic	1733.9	1733.4	1887.9	1887.8
Log-normal	1713.5	1713.2	1860.7	1860.6
Exponential	1769.3	1768.9	1942.0	1941.6
Gompertz	1772.9	1775.0	1943.5	1943.5

Abbreviations: ITC, indirect treatment comparison; OS, overall survival; PFS, progression-free survival

Fractional polynomials

Survival curves based on the 1st and 2nd order fractional polynomial method are presented for OS and PFS in Appendix D (ITC results, supplementary data). Model fit statistics for both OS and PFS (PFS) are presented in Table 24.

Table 24 Fit statistics - Fractional polynomial method

OS				**PFS**			
1st order fractional polynomial		**2nd order fractional polynomial**		**1st order fractional polynomial**		**2nd order fractional polynomial**	
Model	**DIC**	**Model**	**DIC**	**Model**	**DIC**	**Model**	**DIC**
P=-1	1722.8	P1=-0.5, P2=0	1716.5	P=-1	1910.4	P1=-0.5, P2=0	1852.1
P=-0.5	1739.5	P1=-1, P2=0	1713.9	P=-0.5	1932.0	P1=-1, P2=0	1840.3
P=0	1757.7	P1=-1, P2=-1	1711.9	P=0	1945.9	P1=-1, P2=-1	1825.0
P=0.5	1769.0	P1=-1, P2=0.5	1716.2	P=0.5	1947.6	P1=-1, P2=0.5	1850.4
P=1	1773.0	P1=-1, P2=1	1718.3	P=1	1943.6	P1=-1, P2=1	1858.1

Abbreviations: DIC, deviance information criteria; OS, overall survival; PFS, progression-free survival

Uncertainties in the indirect and mixed treatment comparisons

The current study utilised a robust methodology for the SLR with a systematic and comprehensive review and collection of data. However, some limitations apply. Unlike the CABOSUN study, the COMPARZ study included patients with a favourable prognostic risk profile and Kaplan-Meier data results were not available for intermediate and poor prognostic risk groups separately from the ITT population. To account for this, RE and FE models were run and additional fractional polynomial models were tested for statistical fit and clinical plausibility. To further explore the impact of differences in the subgroup data, an additional analysis was carried out on hazard ratios (HRs). Unlike Kaplan-Meier data, HRs were available by subgroup and these were compared despite the violation of the proportional hazards assumption (see Appendix D for further details). This analysis showed that the results for intermediate and poor risk groups are consistent with the results for the overall population.

The data were systematically extracted from the identified studies and the quality of evidence for the analysis was assessed and appraised using the NICE checklist.[68] There are several potential sources of bias with respect to the SLR, which should be noted. There is a potential for bias in the subjective review process, though use of two independent reviewers in the SLR was employed to mitigate this bias potential. Additionally, the SLR may have been impacted by the decision to exclude certain studies due to data availability. Although the NICE checklist was used to assess the quality of evidence, this grading process is subject to bias from the reviewer.

Choice of model

Details on the choice of model and the parameter estimates used to inform the economic model are provided in Section B.3.

The ITC results suggest that cabozantinib significantly increases PFS compared with sunitinib and pazopanib. Overall survival was longest for cabozanitib.

B.2.10 Adverse reactions

No other studies were identified that report additional adverse effects to those reported below.

Summary

- The overall incidence of all-causality AEs in CABOSUN was generally similar between the cabozantinib and sunitinib arms, and the adverse events were similar to those observed with other VEGFR-TKIs in RCC.[44]
- The proportion of patients with grade 3 or 4 AEs was also similar between cabozantinib and sunitinib (68% and 65%, respectively).[44,47]
- The proportion of patients discontinuing study treatment due to AEs was similar for each treatment (21% and 22% for cabozantinib and sunitinib, respectively).[44,47]

Drug exposure and dose intensity

There was a longer duration of exposure in the cabozantinib arm compared with the sunitinib arm in CABOSUN (median: 6.5 months vs. 3.1 months).[44,47] Dose reductions (46% of patients with cabozantinib, 35% with sunitinib) and dose interruptions (73% and 71%, respectively) were frequent with both treatments. The median average daily dose was 50.3 mg in the cabozantinib arm and 44.7 mg in the sunitinib arm. Median dose intensity was 84% for the cabozantinib arm and 89% for the sunitinib arm.[44]

Adverse events

Adverse events (AEs) in CABOSUN were defined as solicited or unsolicited for the purposes of reporting.

- Solicited: eleven events were considered 'expected' per the protocol and presence/ absence and severity were solicited at baseline and for each cycle of treatment: Alanine aminotransferase (ALT) increased, Aspartate aminotransferase (AST) increased, blood bilirubin increased, ECG QT prolonged, fatigue, hypertension, neutrophil count decreased, palmar-plantar erythrodysaesthesia syndrome (PPES), platelet count decreased, diarrhoea, and pancreatitis. [36]
- Unsolicited: other AEs were collected as unsolicited events.[36]

For unsolicited AEs, the Investigator recorded Grade 1-2 AEs that were regarded as related and Grade 3-5 AEs regardless of attribution. For solicited AEs, the Investigator recorded all available CTCAE grades.[36]

The safety data reported here are taken from the CSR[44] and may differ from published data[2] due to regulatory reporting requirements.

The overall incidence of all-causality AEs in CABOSUN was generally similar between the treatment arms. The AEs reported with cabozantinib were consistent with those observed with other VEGFR-TKIs in RCC, and also with its known safety profile in previously treated advanced RCC. AEs are summarised in Table 25.

Table 25 Summary of AE incidence (safety population)

	Cabozantinib N = 78 N (%)	**Sunitinib N = 72 n (%)**
AE	75 (96)	71 (99)
Related AE	74 (95)	70 (97)
Worst AE, grade 3 or 4	53 (68)	47 (65)
Worst related AE, grade 3 or 4	47 (60)	45 (63)
Grade 5 AE up to 30 days after last dose of study treatment[a]	3 (3.8)	6 (8.3)
Grade 5 AE > 30 days after last dose of study treatment	1 (1.3)	3 (4.2)
Related grade 5 AE at any time	2 (2.6)	4 (5.6)
Serious AE	38 (49)	37 (51)
Related serious AE[b]	28 (36)	26 (36)
Deaths	38 (49)	43 (60)
Death up to 30 days after last dose of study treatment	4 (5.1)	8 (11)
Death > 30 days after last dose of study treatment	34 (44)	35 (49)
Discontinuation of study due to AE[c]	21%	22%

a, Grade 5 AEs were not reported for 3 subjects (1 cabozantinib, 2 sunitinib) who died < 30 days after the last dose of study treatment; b, grade 1 or 2 SAEs that did not entail hospitalisation ≥ 24 h were not recorded in the clinical database; c, based on patient disposition, not excluding events of disease progression, only % reported. 'Unsolicited' grade 1 and 2 events not related to study treatment were not collected. Abbreviations: AE, adverse event. Source: Clinical Study Report[44]

- The overall incidence of treatment-related AEs was 95% in the cabozantinib arm and 97% in the sunitinib arm.
- The most frequent treatment-related AEs of any grade (≥ 20% of subjects; solicited or unsolicited) in the cabozantinib arm, in descending order of frequency, were diarrhoea, fatigue, AST increased, hypertension, ALT increased, decreased appetite, PPES, dysgeusia, platelet count decreased, stomatitis, nausea, weight decreased, anaemia, dyspepsia, hypophosphatemia, hypothyroidism, blood creatinine increased, dysphonia, and hypomagnesemia.
- Grade 3/4 all-causality AEs were reported for 68% of patients in the cabozantinib arm and 65% in the sunitinib arm.
- The most frequent treatment-related Grade 3/4 AEs (≥ 5% of patients) in the cabozantinib arm were hypertension (22.0%), diarrhoea (9.0%), hypophosphatemia (9.0%), palmar-plantar erythrodysaesthesia syndrome (7.7%), fatigue (5.1%), decreased appetite (5.1%), and stomatitis (5.1%).

- Grade 5 AEs with death dates up to 30 days after the last dose of study treatment were reported for three patients (3.8%) in the cabozantinib arm and six patients (8.3%) in the sunitinib arm.
- The proportion of patients discontinuing study treatment due to AEs was similar for each treatment (21% and 22% for cabozantinib and sunitinib, respectively).[44,47]

A detailed tabulation of treatment-related AEs in CABOSUN is provided in Table 26.

Table 26 Frequent treatment-related adverse events by grade, safety population

	Cabozantinib N = 78 n (%)		**Sunitinib N = 72 n (%)**	
	Grade		**Grade**	
Preferred term	**Any**	**3/4**	**Any**	**3/4**
Number of subjects with at least one related AE	74 (95)	47 (60)	70 (97)	45 (63)
Solicited related AEs				
Number of subjects with at least one solicited related AE	73 (94)	32 (41)	68 (94)	38 (53)
Diarrhoea	56 (72)	7 (9.0)	35 (49)	6 (8.3)
Fatigue	48 (62)	4 (5.1)	48 (67)	12 (17)
Aspartate aminotransferase (AST) increased	47 (60)	1 (1.3)	18 (25)	2 (2.8)
Hypertension	44 (56)	17 (22)	27 (38)	13 (18)
Alanine aminotransferase (ALT) increased	42 (54)	3 (3.8)	18 (25)	0
Palmar-plantar erythrodysaesthesia syndrome	33 (42)	6 (7.7)	23 (32)	3 (4.2)
Platelet count decreased	29 (37)	1 (1.3)	42 (58)	8 (11)
Neutrophil count decreased	12 (15)	0	25 (35)	3 (4.2)
Blood bilirubin increased	11 (14)	0	4 (5.6)	0
Electrocardiogram QT prolonged	2 (2.6)	0	5 (6.9)	2 (2.8)
Pancreatitis	1 (1.3)	0	0	0
Unsolicited related AEs (reported in ≥ 10% of patients in either treatment arm)				
Number of subjects with at least one unsolicited related AE	69 (88)	33 (42)	62 (86)	30 (42)
Decreased appetite	35 (45)	4 (5.1)	22 (31)	1 (1.4)
Dysgeusia	32 (41)	0	21 (29)	0
Stomatitis	29 (37)	4 (5.1)	21 (29)	4 (5.6)
Nausea	24 (31)	2 (2.6)	26 (36)	2 (2.8)
Weight decreased	24 (31)	3 (3.8)	12 (17)	0
Anaemia	23 (29)	0	32 (44)	1 (1.4)

Dyspepsia	18 (23)	0	12 (17)	0
Hypophosphataemia	18 (23)	7 (9.0)	12 (17)	5 (6.9)
Hypothyroidism	17 (22)	0	4 (5.6)	0
Blood creatinine increased	16 (21)	1 (1.3)	14 (19)	2 (2.8)
Dysphonia	16 (21)	1 (1.3)	1 (1.4)	0
Hypomagnesaemia	16 (21)	1 (1.3)	8 (11)	0
Dry mouth	15 (19)	0	8 (11)	0
Dry skin	15 (19)	0	6 (8.3)	0
Vomiting	15 (19)	1 (1.3)	15 (21)	1 (1.4)
Constipation	14 (18)	1 (1.3)	9 (13)	0
Dizziness	14 (18)	0	12 (17)	0
Alopecia	13 (17)	0	2 (2.8)	0
Hypoalbuminaemia	13 (17)	0	12 (17)	0
Hypocalcaemia	13 (17)	2 (2.6)	10 (14)	0
Dermatitis acneiform	12 (15)	0	2 (2.8)	0
Dyspnoea	12 (15)	1 (1.3)	8 (11)	1 (1.4)
Hypokalaemia	12 (15)	1 (1.3)	4 (5.6)	0
Rash maculo-papular	11 (14)	0	9 (13)	2 (2.8)
Blood alkaline phosphatase increased	10 (13)	0	8 (11)	0
Lymphocyte count decreased	10 (13)	0	13 (18)	3 (4.2)
Headache	9 (12)	0	11 (15)	1 (1.4)
White blood cell count decreased	9 (12)	0	25 (35)	2 (2.8)
Abdominal pain	8 (10)	0	5 (6.9)	2 (2.8)
Dehydration	8 (10)	3 (3.8)	7 (9.7)	1 (1.4)
Hyperglycaemia	8 (10)	0	8 (11)	2 (2.8)
Insomnia	8 (10)	0	6 (8.3)	0
Peripheral sensory neuropathy	8 (10)	1 (1.3)	3 (4.2)	0
Hyponatraemia	6 (7.7)	2 (2.6)	13 (18)	4 (5.6)
Oedema peripheral	5 (6.4)	0	9 (13)	0
Proteinuria	5 (6.4)	2 (2.6)	9 (13)	1 (1.4)
Muscular weakness	3 (3.8)	0	10 (14)	0

Denominators for percentages are N, the total number of patients in each treatment arm. At each level of summarisation, a patient was counted once for the most severe event if the patient reported one or more events. Note that there is no CTCAE Grade 1 category for pancreatitis or Grade 4 category for fatigue or PPES. Abbreviations: AE, adverse event. Source: Clinical Study Report [44]

B.2.11 Ongoing studies

None.

B.2.12 Innovation

The superior efficacy results of cabozantinib compared with current treatments seen both in CABOSUN and in the METEOR trial of previously treated advanced RCC[69,70] may 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 (MET, AXL), as well as tumour angiogenesis (VEGF).[71] By targeting MET and AXL receptors in addition to VEGFR, cabozantinib may provide additional anticancer efficacy over the more selective, existing anti-VEGFR agents.

Cabozantinib was granted Promising Innovative Medicine (PIM) designation under the Early Access to Medicines Scheme (EAMS) in July 2016.

B.2.13 Interpretation of clinical effectiveness and safety evidence

Conclusions on clinical effectiveness

The clinical benefit of cabozantinib in previously untreated patients with advanced RCC was demonstrated by the statistically and clinically significant prolongation of IRC-assessed PFS compared with sunitinib: stratified HR was 0.48 (95% CI 0.31, 0.74; 2-sided p-value = 0.0008), corresponding to a 52% reduction in the risk of disease progression or death; median PFS was 8.6 months in the cabozantinib arm and 5.3 months in the sunitinib arm, an estimated 3.3-month difference in the medians.[44,47]

Cabozantinib was also associated with a numerical improvement in OS compared with sunitinib: median OS was 30.3 months in the cabozantinib arm vs. 21.0 months in the sunitinib arm, an estimated 9.3-month difference (HR 0.74 [95% CI 0.47, 1.14]; 2-sided p-value = 0.1700).[44] While the survival data are still immature, the latest OS results[47] were consistent with those previously reported.

CABOSUN was not conducted as a registration trial, and consequently there are some limitations to the data; these were discussed by Rini et al[72] in relation to the original study publication. Many of the original concerns (non-standard censoring rules, non-blinded investigator assessment of response and progression) were

addressed by the IRC reassessment and FDA-guided censoring analyses carried out for the regulatory submission, on which the current submission to NICE is based.

There was an imbalance in the number of patients with missing data. One patient in the cabozantinib arm and six in the sunitinib arm withdrew from the study prior to receiving study treatment. The reasons for these withdrawals are not known.[44] There was also a higher incidence of missing or unevaluable data in the sunitinib arm. Six patients in the cabozantinib arm and 18 in the sunitinib arm were not evaluable because they had no adequate post-baseline imaging assessments: cabozantinib: adverse event (5), withdrew consent (1); sunitinib: adverse event (6), death (2), disease progression (1), withdrew consent (9).[47] Because of the nature of these clinical events, none of these patients was likely to have experienced a response or prolonged PFS.[73] For the CSR analyses, patients with missing data were censored for PFS according to FDA guidance (see Section B.2.6, Table 12). Based on their baseline characteristics, the sunitinib patients without post-baseline imaging would not be expected to have a better prognosis than sunitinib patients who had a response recorded, and therefore it is unlikely that the radiographic endpoints were biased against sunitinib by these missing data.

Rini et al also suggested that PFS with sunitinib in CABOSUN was lower than expected. However, data highlight that this is not the case:

- The registry analysis by Ko et al.[10] of outcomes for 1189 previously untreated poor and intermediate risk patients receiving targeted therapies (among whom sunitinib was the most common treatment), found a PFS of 5.6 months, which is consistent with the result in CABOSUN.
- In addition, as explained by Choueiri et al.,[73] an important distinction of the CABOSUN study was the focus on a population of patients with high rates of poor prognostic clinical factors, which set it apart from other contemporary trials that involved untreated patients with metastatic RCC, such as COMPARZ[27] (pazopanib vs. sunitinib) and RECORD3[74] (sunitinib vs. everolimus). The differences between these populations are shown in Table 27. COMPARZ and RECORD3 had significant proportions of favourable-risk patients, whereas

favourable-risk patients were not eligible for CABOSUN. CABOSUN also had a larger proportion of patients with bone metastases than the other trials.

- A retrospective analysis of the registration Phase 3 sunitinib vs. IFN-α trial in treatment-naïve subjects (Rini et al 2017)[75] also showed a substantial difference in median PFS in the sunitinib arm by IMDC risk category (favourable 16.0 months; intermediate 10.7 months; poor 2.5 months).
- Choueiri et al.[73] also note that, in an all-comer population, PFS with sunitinib has been found to be generally shorter in the community setting (Schnadig et al. 2014,[76] PFS = 7.5 months). Choueiri et al. state that 'the cooperative group setting is more reflective of community practice'.[73]

Table 27 Selected prognostic factors in contemporary trials in advanced RCC

Study	Risk model	Favourable risk group (%)	Bone metastases (%)	ECOG PS2 (%)
CABOSUN	IMDC	0	36	13
RECORD3	MSKCC	30	23	6
COMPARZ	IMDC	25	18	NR

Abbreviations: ECOG, Eastern Cooperative Oncology Group; IMDC, International Metastatic Database Consortium Criteria; MSKCC, Memorial Sloan Kettering Cancer Center; NR, not reported; PS, performance score. Source: Choueiri et al 2017[73]

Conclusions on safety

Results from CABOSUN add to the previous experience with cabozantinib from the METEOR study, which led to approval for cabozantinib in patients with advanced RCC who had received prior anti-angiogenic therapy, and established the safety profile for cabozantinib. The AE profile of cabozantinib was similar across both studies, with no new safety signals identified in CABOSUN.[44] The safety profile of cabozantinib is similar to that of sunitinib and other established TKIs. In general, AEs from cabozantinib can be managed with supportive care and dose reductions.[1]

End-of-life criteria

Cabozantinib for untreated locally advanced or metastatic RCC meets the end-of-life criteria. See Table 28 for details.

Table 28 End-of-life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	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.[11]	Section B.1.3, page 13
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	In the CABOSUN trial,[44] median OS 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 (i.e. 7.9 months) compared with sunitinib in the base case pairwise comparison. The other treatment currently recommended by NICE and 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[27]). In the economic modelling, cabozantinib is associated with a gain of 0.80 life years (9.6 months) compared with pazopanib	Trial: Section B.2.6, page 40 Economic modelling: Section B.3.7, p. 123

Abbreviations: CI, confidence interval; IMDC, International Metastatic RCC Database Consortium; NE, not estimable; OS, overall survival; PFS, progression-free survival; RCC, renal cell carcinoma; VEGF, vascular endothelial growth factor; NHS, National Health Service

B.3 Cost effectiveness

B.3.1 Published cost-effectiveness studies

Details of the search, the identified studies, and quality assessments are provided in Appendix G.

The search identified 23 published cost-effectiveness studies, of which seven were conducted from an English, Welsh or British perspective. These are summarised in Table 29. The details of the search, all the identified studies, and quality assessments are reported in Appendix G.

Table 29 Summary list of published cost-effectiveness studies

Study	Year of costing	Country	Aim	Patient population	Summary of model	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Amdahl 2017[77]	2014	UK	To provide a direct comparison of the cost-effectiveness of pazopanib vs. sunitinib from the perspective of the UK NHS based on data from COMPARZ and other sources	mRCC	Partitioned-survival analysis model with three health states (pre-progression, post-progression, or death); time horizon 5 years (duration of follow-up for final survival analysis in COMPARZ); treatment cycle length was one week in order to accommodate the 4-week cycle for pazopanib and the 6-week cycle for sunitinib.	1.6026 (pazopanib); 1.5432 (sunitinib); 0.0595 (incremental); discounted annually at 3.5%	£38,126 (pazopanib); £39,038 (sunitinib); –£912 (incremental); discounted annually at 3.5%	NR; pazopanib is dominant
Hoyle 2010[78]	2007/2008	UK	To estimate the cost-effectiveness of temsirolimus vs. IFN- α for first-line treatment of patients with advanced, poor prognosis RCC, from the perspective of the UK NHS	Phase III trial of temsirolimus, aRCC	Decision-analytic model, using techniques from survival analysis to consider progression of RCC in a cohort of patients over time; subsequent second-line treatments, such as sunitinib and sorafenib, are not explicitly modelled; at any time, a patient is assumed to be in one of three health states:	0.77 (temsirolimus); 0.53 (IFN- α); 0.24 (incremental temsirolimus vs. IFN- α); discounted annually at 3.5%	£28,849 (temsirolimus); £6,519 (IFN- α); £22,331 (incremental temsirolimus vs. IFN- α); discounted annually at 3.5%	£94,632 per QALY gained (temsirolimus vs. IFN- α, general study population); £150,721 per QALY gained (temsirolimus vs. IFN- α, clear-cell population); £49,701 per LY gained (temsirolimus vs. IFN- α, general study population); £80,229 per LY gained (temsirolimus vs. IFN- α,

					PFS, progressive disease and death. The model uses a 10-year time horizon (effectively lifetime because virtually all simulated patients are dead at 10 years), and a 6-week model cycle.			clear-cell population)
HTA 2007[79]	NR	Wales	To provide recommendation for sunitinib for treatment of aRCC and/or mRCC - not recommended for use within NHS Wales	1L mRCC	Manufacturer's submission of a cost-utility evaluation to compare sunitinib vs. IFN- α as first-line treatment for patients with aRCC and/or mRCC. IFN- α represented the current standard treatment for this patient group in Wales. The model did not address the use of sunitinib as a second-line therapy. The decision analytic model was populated with data derived from trial A6181034. The study enrolled adults with previously untreated aRCC and/or mRCC with an ECOG performance score of 0 or 1 and the analysis was carried out over the remaining	0.69 (incremental sunitinib vs. IFN-α); discounted annually at 3.5%	£23,366 per patient per year (sunitinib); £6,634 per patient per year (IFN- α); £20,283 (incremental sunitinib vs. IFN-α); discounted annually at 3.5%	£29,199 per QALY gained (£36,225 for 5 years horizon); £24,801 per LYG; £53,909 per PFYG (sunitinib vs. IFN- α)

					life expectancy of the individual patient, with a time horizon of six years.			
HTA 2009[9]	NR	England, Wales	To provide guidance for bevacizumab (first-line), sorafenib (first and second-line), sunitinib (second-line) and temsirolimus (first-line) for the treatment of aRCC or mRCC	1L aRCC/mRCC; 9% of patients receiving bevacizumab plus IFN- α and 8% of patients receiving IFN- α plus placebo were defined as having a poor prognosis	Manufacturers' models: Bevacizumab: simple state-transition model with three health states: PFS, progressive disease, death. The model compared bevacizumab plus IFN- α with IFN- α plus placebo as a first-line treatment for people suitable for immunotherapy. Temsirolimus: state-transition model with three health states: PFS, post-progression and death. The PFS state was then subdivided into stable disease, complete/partial response and progressive disease. The model compared temsirolimus with IFN- α as first-line treatment for people with at least three of six risk factors for poor prognosis, who were suitable for	NR; discounted annually at 3.5%	NR; discounted annually at 3.5%	£74,999 (manufacturer's model, bevacizumab plus IFN- α vs. IFN- α plus placebo, people suitable for immunotherapy); £55,814 (manufacturer's model, temsirolimus vs. IFN- α, people with poor prognosis suitable for immunotherapy); £57,731 (manufacturer's model, temsirolimus vs. IFN- α, people with poor prognosis suitable for immunotherapy, subgroup with clear cell carcinoma); £51,159 (manufacturer's model, temsirolimus vs. IFN- α, people with poor prognosis suitable for immunotherapy, subgroup with non-clear-cell carcinoma); £81,201 (manufacturer's model, temsirolimus vs. BSC, patients with poor prognosis unsuitable for immunotherapy); £171,301 (assessment group model,

					immunotherapy. Sorafenib: simple state-transition model with three health states: PFS, progressed disease and death. The model compared sorafenib with BSC for people in whom immunotherapy had failed or who were unsuitable for immunotherapy. Assessment group model: Markov model with three treatment strategy questions: first-line treatment (bevacizumab plus IFN- α compared with IFN- α); first-line treatment of people with a poor prognosis (temsirolimus compared with IFN- α) and second-line treatment (sorafenib compared with BSC) using similar model structures but with different model parameter data for each question. The model used three distinct health states: PFS, progressive			bevacizumab plus IFN- α vs. IFN- α); £94,385 (assessment group model, temsirolimus vs. IFN- α, people with poor prognosis)

					disease and death.			
HTA 2009[19]	NR	UK	To provide guidance for sunitinib for first-line treatment of aRCC and/or mRCC	1L aRCC/mRCC	Manufacturers' models: Sunitinib: simple state-transition model with three health states: PFS, progressed disease, death. The model compared sunitinib with IFN- α as a first-line treatment for people suitable for immunotherapy. Assessment group model: Markov model with three distinct health states: PFS, progressive disease, death, to estimate the cost-effectiveness of sunitinib, sorafenib, temsirolimus and bevacizumab plus IFN-α, against relevant comparators and according to the licensed indication of each drug.	NR; discounted annually at 3.5%	NR; discounted annually at 3.5%	£72,003 (manufacturer's model sunitinib vs. IFN-α); £71,760 (manufacturer's model, population censored for crossover); £41,472 (manufacturer's model, participants who received no systemic post-study treatments); £104,715 (assessment group model); £62,365 (assessment group model, participants who received no systemic post-study treatments)
Kilonzo 2013[80]	2010	UK	To provide a description of the company submission, the ERG review and NICE's subsequent	1L aRCC and/or mRCC	Partitioned survival model with three mutually exclusive health states: alive pre-progression, alive post-progression and dead; using parametric	Pazopanib: 1.071 (ITT); 1.616 (Cox model censored on crossover on receipt of other	Pazopanib: £32,099 (ITT); £34,676 (Cox model censored on crossover on receipt of other	Sunitinib vs. pazopanib: £1,790 (base case: RPSFT weighted unadjusted); Other approaches: £4,936 (ITT; sunitinib

			decisions for pazopanib for the first-line treatment of patients with aRCC and/or mRCC.		survival curves fitted to empirical data on OS and PFS over time.	anti-cancer therapy); 1.613 (IPCW); 1.966 (RPSFT weighted unadjusted); 2.697 (RPSFT unweighted adjusted); 2.385 (no post-study therapy) -0.068 (incremental sunitinib vs. pazopanib; pazopanib more effective); -0.717 (incremental, IFN vs. pazopanib; pazopanib more effective); -0.979 (incremental, BSC vs. pazopanib; pazopanib more effective)	anti-cancer therapy); £34,661 (IPCW); £36,301 (RPSFT weighted unadjusted); £39,689 (RPSFT unweighted adjusted); £38,241 (no post-study therapy) £-122 (incremental, sunitinib vs. pazopanib; pazopanib more costly); £-27,900 (incremental, IFN vs. pazopanib; pazopanib more costly); £-32,200 (incremental, BSC vs. pazopanib; pazopanib more costly)	more effective and more costly) £5,327 (Cox model censored on crossover on receipt of other anti-cancer therapy, sunitinib more effective and more costly) £5,139 (IPCW; sunitinib more effective and more costly) £4,394 (RPSFT; unweighted adjusted); £4,238 (no post-study therapy) IFN- α vs. pazopanib: £38,925 (base case: RPSFT weighted unadjusted); Other approaches: ITT: dominated by pazopanib £71,648 (Cox model censored on crossover on receipt of other anti-cancer therapy); £72,274 (IPCW); £21,625 (RPSFT, unweighted adjusted); £26,293 (no post-study therapy)

								BSC vs. pazopanib: £32,898 (base case: RPSFT weighted unadjusted); Other approaches: £322,237 (ITT); £48,638 (Cox model censored on crossover on receipt of other anti-cancer therapy); £48,877 (IPCW); £20,824 (RPSFT, unweighted adjusted); £24,438 (no post-study therapy)
Thompson Coon 2010[81]	NR	UK	To assess the cost-effectiveness of sunitinib, sorafenib, bevacizumab plus IFN, and temsirolimus against relevant comparators for licensed indications.	1L aRCC (sunitinib, bevacizumab plus IFN- α); 1L aRCC and poor prognosis (temsirolimus); 1L aRCC/mRCC RCC (sunitinib, bevacizumab plus IFN- α); 1L aRCC/mRCC RCC and poor prognosis	Manufacturers' models: Sunitinib: Pfizer manufacturer model for sunitinib vs. IFN; Pfizer model with PenTAG adjustment, for sunitinib vs. IFN; Pfizer manufacturer model, bevacizumab plus IFN vs. IFN; Roche manufacturer model, bevacizumab plus IFN vs. IFN; Roche model with PenTAG adjustment, bevacizumab plus IFN vs. IFN;	0.60 (Pfizer manufacturer model, incremental, sunitinib vs. IFN); 1.65 (Pfizer manufacturer model, bevacizumab + IFN); 1.31 (Pfizer manufacturer model, IFN); 0.34 (Pfizer manufacturer model, incremental, bevacizumab +	- £54,984 (Pfizer manufacturer model, bevacizumab + IFN); £18,001 (Pfizer manufacturer model, IFN); £36,923 (Pfizer manufacturer model, incremental, bevacizumab +	£28,546 per QALY, £21,116 per LYG (Pfizer manufacturer model, sunitinib vs. IFN); £48,052 Pfizer model with PenTAG adjustment1; £107,357 per QALY, £81,754 per LYG (Pfizer manufacturer model, bevacizumab + IFN vs. IFN); £75,000 per QALY, £58,712 per LYG (Roche manufacturer model, bevacizumab + IFN vs. IFN);

				(temsirolimus)	Wyeth manufacturer model, temsirolimus vs. IFN; Evidence Review Group (PenTAG) cost-utility model: Decision-analytic Markov-type model to simulate disease progression and estimate cost-effectiveness of the drugs under consideration; three health states: PFS, progressive disease, and death. Weibull survival curves were fitted to the PFS and OS KM curves from clinical trials for the baseline comparator. Relative measures of treatment effectiveness (HRs) were then used to estimate the expected disease progression compared with baseline.	IFN vs. IFN); 0.27 (Roche manufacturer model, incremental, bevacizumab + IFN vs. IFN); 0.13 (Wyeth manufacturer model, incremental, temsirolimus vs. IFN); 0.50 (Wyeth manufacturer model, temsirolimus, clear cell patient subgroup); 0.39 (Wyeth manufacturer model, IFN, clear cell patient subgroup); 0.109 (Wyeth manufacturer model, incremental, temsirolimus vs. IFN, clear cell patient subgroup); 0.55 (Wyeth	IFN vs. IFN); £20,000 (Roche manufacturer model, incremental, bevacizumab + IFN vs. IFN); £7,493 (Wyeth manufacturer model, incremental, temsirolimus vs. IFN); £33,429 (Wyeth manufacturer model, temsirolimus, clear cell patient subgroup); £27,139 (Wyeth manufacturer model, IFN, clear cell patient subgroup); £6,291 (Wyeth manufacturer model, incremental, temsirolimus vs. IFN, clear cell patient subgroup); £34,601 (Wyeth	£117,000 (Roche model with PenTAG adjustment, bevacizumab plus IFN vs IFN); £55,814 per QALY, £35,577 per LYG (Wyeth manufacturer model, temsirolimus vs. IFN); £57,731 per QALY, £39,188 per LYG (Wyeth manufacturer model, temsirolimus vs. IFN, clear cell patient subgroup); £51,159 per QALY, £29,035 per LYG (Wyeth manufacturer model, temsirolimus vs. IFN, non-clear cell patient subgroup); £102,000 (Wyeth model with PenTAG adjustment; resource use1, temsirolimus vs IFN); £121,300 (Wyeth model with PenTAG adjustment; resource

						manufacturer model, temsirolimus, non-clear cell patient subgroup); 0.29 (Wyeth manufacturer model, IFN, non-clear cell patient subgroup); 0.260 (Wyeth manufacturer model, incremental, temsirolimus vs. IFN, non-clear cell patient subgroup); 0.51 (Wyeth manufacturer model, temsirolimus); 0.30 (Wyeth manufacturer model, BSC); 0.205 (Wyeth manufacturer model, incremental, temsirolimus vs. BSC); 1.62 (PenTAG model, sunitinib);	manufacturer model, temsirolimus, non-clear cell patient subgroup); £21,296 (Wyeth manufacturer model, IFN, non-clear cell patient subgroup); £13,305 (Wyeth manufacturer model, incremental, temsirolimus vs. IFN, non-clear cell patient subgroup); £33,612 (Wyeth manufacturer model, temsirolimus); £16,932 (Wyeth manufacturer model, BSC); £16,680 (Wyeth manufacturer model, incremental, temsirolimus vs. BSC); £39,623 (PenTAG model, sunitinib);	use1, temsirolimus vs IFN, clear cell patient subgroup); £63,100 (Wyeth model with PenTAG adjustment; resource use2, temsirolimus vs IFN, non-clear cell patient subgroup); £81,201 per QALY, £43,746 per LYG (Wyeth manufacturer model, temsirolimus vs. BSC); £71,462 per QALY, £58,647 per LYG (PenTAG model, sunitinib vs. IFN); £171,301 per QALY, £133,952 per LYG (PenTAG model, bevacizumab + IFN vs. IFN); £81,687 per QALY, £42,902 per LYG (PenTAG model, temsirolimus vs. IFN, poor prognosis); £128,872 per QALY,

						1.19 (PenTAG model, IFN); 0.44 (PenTAG model, incremental, sunitinib vs. IFN); 1.45 (PenTAG model, bevacizumab + IFN); 1.19 (PenTAG model, IFN); 0.27 (PenTAG model, incremental, bevacizumab + IFN vs. IFN); 0.77 (PenTAG model, temsirolimus, poor prognosis); 0.53 (PenTAG model, IFN, poor prognosis); 0.24 (PenTAG model, incremental, temsirolimus vs. IFN, poor prognosis); 0.65 (PenTAG model,	£8,438 (PenTAG model, IFN); £31,185 (PenTAG model, incremental, sunitinib vs. IFN); £53,873 (PenTAG model, bevacizumab + IFN); £8,438 (PenTAG model, IFN); £45,435 (PenTAG model, incremental, bevacizumab + IFN vs. IFN); £25,794 (PenTAG model, temsirolimus, poor prognosis); £6,519 (PenTAG model, IFN, poor prognosis); £19,276 (PenTAG model, incremental, temsirolimus vs. IFN, poor prognosis); £21,256 (PenTAG model,	£68,599 per LYG (PenTAG model, temsirolimus vs. IFN, poor prognosis, clear cell patient subgroup); £89,394 per QALY, £58,378 per LYG (PenTAG model, temsirolimus vs. IFN, poor prognosis, non-clear cell patient subgroup)

						temsirolimus, poor prognosis, clear cell patient subgroup); 0.53 (PenTAG model, IFN, poor prognosis, clear cell patient subgroup); 0.11 (PenTAG model, incremental, temsirolimus vs. IFN, poor prognosis, clear cell patient subgroup); 1.17 (PenTAG model, temsirolimus, poor prognosis, non-clear cell patient subgroup); 0.53 (PenTAG model, IFN, poor prognosis, non-clear cell patient subgroup); 0.64 (PenTAG model, incremental, temsirolimus vs. IFN, poor prognosis, non-clear cell patient	temsirolimus, poor prognosis, clear cell patient subgroup); £6,519 (PenTAG model, IFN, poor prognosis, clear cell patient subgroup); £14,737 (PenTAG model, incremental, temsirolimus vs. IFN, poor prognosis, clear cell patient subgroup); £63,418 (PenTAG model, temsirolimus, poor prognosis, non-clear cell patient subgroup); £6,519 (PenTAG model, IFN, poor prognosis, non-clear cell patient subgroup); £56,899 (PenTAG model, incremental, temsirolimus vs. IFN, poor prognosis, non-clear cell patient	

						subgroup); discounted annually at 3.5%	subgroup); discounted annually at 3.5%	

Abbreviations: 1L, first line; aRCC, advanced renal cell carcinoma; BSC, best supportive care; CAN$, Canadian Dollar; COMPARZ, Comparing the Efficacy, Safety and Tolerability of Pazopanib versus Sunitinib (trial); ECOG, Eastern Cooperative Oncology Group; GBP, Great Britain Pound; HR, Hazard ratio; ICER, Incremental cost-effectiveness ratio; IFN, Interferon;; IRC, Independent Review Committee; ITT, Intent-to-treat (population); LYG, life-years gained; MIU, million international units; mRCC, metastatic renal cell carcinoma; MSKCC, Memorial Sloan Kettering Cancer Centre; NHS, National Health Service; NR, Not Reported; OS, Overall survival; PenTAG, Peninsula Technology Assessment Group; PFLYG, progression-free life-years gained; PFS, Progression-free survival; QALY, quality-adjusted life year; RCC, renal cell carcinoma; RCT, Randomised controlled trial; RPSFT, Rank preserving structural failure time

1 Weibull curve is fitted to fewer data points, in this case one data point per month

2 Resource based on clinical opinion and assumption that 25% of patients have IFN administered by a district nurse, at a cost of £25 per visit, and the remaining 75% self-inject, at no cost

B.3.2 Economic analysis

The identified models used both partitioned survival and Markov-transition model approaches. The partitioned survival approach is similar to the Markov cohort model because the modelled patients are in one of the three mutually exclusive health states at any given time point. However, unlike the Markov cohort model, the partitioned survival model estimates the proportion of patients from the cohort in each health state based on PFS and OS survival curves. All models included three health states: PFS, progressed disease and death. The model cycle length varied, with the shortest model cycle being one week in the Amdahl et al. 2017 publication (pazopanib versus sunitinib).[77] In that model the treatment cycle length was one week in order to accommodate the 4-week cycle for pazopanib and the 6-week cycle for sunitinib.

No published cost-effectiveness model including cabozantinib, pazopanib and sunitinib was identified, and hence we developed a de novo model comparing these treatments.

Patient population

It is anticipated that cabozantinib will be approved for the treatment of treatment-naive patients with advanced RCC, based on the results of a phase 2 clinical trial (CABOSUN).[44] The current economic analysis evaluates the cost-effectiveness of cabozantinib as first-line treatment for advanced RCC.

The key clinical data source is the CABOSUN study, a randomised, open-label, phase 2 clinical trial comparing cabozantinib with standard-of-care sunitinib in intermediate and poor risk patients with previously untreated advanced RCC.[44] The study is explained in detail in Section B.2. Data from the CABOSUN study were used to inform the cost-effectiveness comparison of cabozantinib versus sunitinib. Comparisons to sunitinib and pazopanib are supported by results from the ITC described in Section B 2.9.

In the CABOSUN study, patients must have been classified as intermediate or poor risk by the IMDC criteria[21] and must not have received prior systemic treatment. The base case of the economic model is based on the previously untreated intermediate

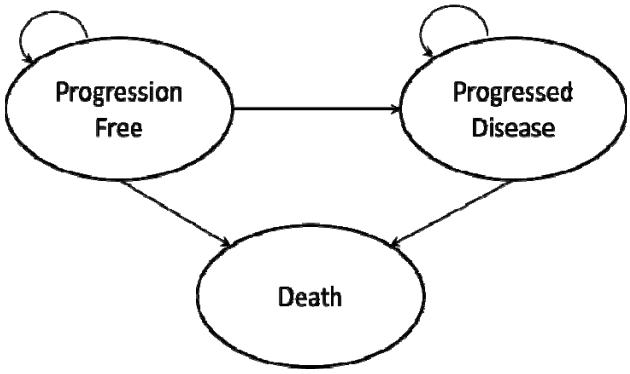
and poor risk population only, whereas the ITC includes studies with mixed populations in terms of baseline prognostic risk group.

Model structure

The cost-effectiveness model was developed in Microsoft Excel® using a partitioned survival model structure in both a deterministic and a probabilistic (Monte Carlo simulation) framework. Unlike a pure Markov approach, which uses explicit transition probabilities for each change in health state, the partitioned survival approach estimates the proportions of patients in each health state based on parametric survival curves fitted to clinical trial data on PFS and OS over time. The structure of the model has been chosen based on identified models for advanced RCC treatment submitted to NICE for completed and ongoing appraisals (GID-TA10123[82], TA215[35], TA333[83], TA219[84]), and opinions from clinical experts. It contains the three mutually exclusive health states that are most relevant from a patient, clinician and NHS perspective:

- Progression-free (PF) – during this stage it is assumed that patients' tumours are in a stable or responding state and not actively progressing. Patients in this state are assumed to incur costs associated with active management, and costs associated with medical management of the condition and grade 3/4 AEs. Patients also experience a higher utility weighting associated with non-progressing disease.
- Progressed disease (PD) – when a patient transitions into the progressed disease health state, primary treatment is terminated and second-line treatment might be initiated within a certain number of weeks. Patients continue to incur costs associated with medical management and palliative care, and experience a lower utility weighting.
- Death – this is an absorbing health state.

Figure 12 Structure of economic model



The model structure is shown in Figure 12. Circles represent health states and the arrows represent transition between states. At any point in time, a patient is assumed to be in one of the three states. Patients move between states at the end of each model cycle. The health states of a cohort of patients are modelled at each discrete model cycle. All patients enter the model in the PF state, where it is assumed that they are treated with the first-line treatment (cabozantinib, sunitinib, or pazopanib). Patients remain in the PF state until they experience disease progression or die. Once patients enter the PD state, first-line treatment is discontinued and some patients are treated with subsequent treatment second-line. Treatment with alternative targeted therapy following progression was the normal strategy in CABOSUN and is thus already captured in the relevant OS survival curves. Patients remain in this state until death.

The model uses estimates of clinical effectiveness, costs and HRQL to model progression of disease and cost-effectiveness over time. A time horizon of 20 years has been chosen to capture the life expectancy of the cohort. The impact of the time horizon on results is explored in a sensitivity analysis. The model applies a cycle length of one week, as this is suitable to capture changes in the health states, allowing for robust estimates of costs and benefits to be calculated for each treatment. Sunitinib is administered daily for 4 weeks followed by 2 weeks without treatment. Pazopanib and cabozantinib are given daily and administered continuously. In addition, a half-cycle correction is implemented to obtain a more accurate estimation of PFS and OS. This model structure is sufficient to capture differences in costs and effects between treatments in patients with advanced RCC. Summary details of the model structure are provided in Table 30. The main inputs for

previous and ongoing NICE technology appraisal economic models in advanced RCC are summarised and compared in

Table 31.

Table 30 Summary of the model structure

	Approach	Source / Justification
Model	Partitioned survival model	It is considered as one of the standard methods for population-based cancer patient survival analysis. Method is in line with previous health economic analyses.[17,35]
Health states	Three health states: progression-free, progressed disease and death.	The model structure and the health states utilised reflect the natural history of the disease. Additionally they are typical of modelling in metastatic oncology and have been utilised in previous NICE STAs.[17,35]
Adverse events	Included in the model as a one-off time event applied during the first cycle. AEs are associated with additional cost and disutility.	Based on observed treatment-emergent grade 3/4 AEs (TEAE) with occurrence in more than 5% of the population in any of the pivotal trials, and judged by clinical expert to have implications for resource use. AEs for pazopanib were obtained from the CSR.[85]
Health related quality of life	Health state specific utility values were obtained from systematic literature review. Before and after progression utilities were assumed to be independent of treatment.	Utilities used in the model were identified in a systematic literature review. All treatments were assumed to have health-state-specific utilities with reductions associated with adverse events experienced by patients.
Resource utilisation and costs	• Treatment costs • AE costs • Disease management costs for PFS health state • Subsequent treatment costs • Disease management costs for progressed health state • Terminal care cost	Based on UK reference costs, literature and expert opinion.
Abbreviations: AE, adverse event; CSR, clinical study report; PFS, progression-free survival; STA, single technology appraisal; TEAE, treatment-emergent adverse event		

Table 31 Features of the economic analysis

	Previous /ongoing appraisals			Current appraisal	
Factors	Tivozanib[1]	Pazopanib[2]	Sunitinb[3]	**Chosen**	**Justification**

				values	
Time horizon	10 years	10 years	10 years	20 years	Long enough to capture all the cost and quality of life benefit.
Comparator	Sunitinib, Pazopanib Interferon (IFN)-α	Sunitinib, Interferon (IFN)-α Best supportive care (BSC)	Interferon (IFN)-α	Sunitinib Pazopanib	In line with the decision problem and scope. Treatment options relevant for UK settings.
Cycle length	1 week	1 day	4 days	1 week	Allow robust estimations of costs and benefits to be calculated for each treatment.
Measurement of health effects	QALYs	QALYs	QALYs	QALYs	NICE reference case[9]
Discount (costs/effects)	3.5% per annum	3.5% per annum	3.5% per annum	3.5% per annum	NICE reference case[9]
Perspective	NHS/PSS	NHS/PSS	NHS/PSS	NHS/PSS	NICE reference case[9]
Model	Partitioned-survival model	Partitioned-survival model	NR	Partitioned-survival model	In line with previous NICE STAs.[1,2,3]
Health states	Three discrete mutually exclusive health states ("Alive pre-progression", "Alive post-progression" and "Dead")	Three discrete mutually exclusive health states ("Alive pre-progression", "Alive post-progression" and "Dead")	Three health states: progression-free survival (PFS), progressive disease (PD) and death	Three health states: progression-free, progressive disease and death	The model structure and the health states utilised reflect the natural history of the disease. Additionally they are typical of modelling in metastatic oncology and have been utilised in previous NICE STAs.[1,2,3]
Adverse events	AE grade 3/4 ≥ 5% incorporated in	AE of particular interest based on clinical	Disutilities of AE were not considered	AE grade 3/4 ≥ 5% incorporated	Those grade 3/4 AEs ≥ 5% are believed to have

	the model. Disutility based on published cost effectiveness analysis of pazopanib[2] based on study VEG105192.[4]	opinion, AE grade 3/4 ≥ 5% and AE all grade ≥20% are considered in the model. Disutility based on VEG105192 CSR. For additional costs associated with AE, only grade 3 and 4 events ≥ 5% were considered.		in the model. Disutility based on published cost effectiveness analysis of pazopanib[2] based on study VEG105192.[4]	important impact on patients' costs and quality of life.
Health related quality of life	Patient data from the TIVO-1 study.[5]	HRQL in progression-free state based on VEG105192 clinical study; HRQL in post-progression state based on Remak 2008[6] and Parasuraman 2008[7]	HRQL data based on the RCT reported by Motzer and colleagues (2007).[8]	HRQL the same as used in Tivozanib NICE STA.[1]	The utility values are thought reasonable based on Tivozanib NICE STA.[1]
Resource utilisation and costs	1.Drug costs 2. AE costs 3. Disease management costs	1.Drug costs 2. AE cost (grade 3+, ≥5%) 3. Disease management costs 4. Costs of progression, and supportive care costs	1. Drug costs 2. An allowance for the mean cost of differences in expected adverse events 3.Disease management costs 4.Costs associated ongoing BSC	1. Drug costs 2. AE cost (grade 3+, ≥5%) 3. Disease management cost 4. End-of-life costs	Acceptable costs applied in the Tivozanib NICE STA.[1]

Abbreviations: PSS, personal social services; QALYs, quality-adjusted life years; NHS, National Health Service; NR, not reported; HRQL, Health related quality of life; AE adverse event; BSC, best supportive care

Sources:

1. Tivozanib for treating renal cell carcinoma [ID591].[17] https://www.nice.org.uk/guidance/indevelopment/gid-ta10123/documents

2. Pazopanib for the first-line treatment of advanced renal cell carcinoma [TA215].[35] https://www.nice.org.uk/guidance/ta215

3. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma [TA169].[19] https://www.nice.org.uk/guidance/ta169

4. Sternberg CN, Hawkins RE, Szczylik C, et al. A randomized, double-blind phase III study (VEG105192) of pazopanib (paz) versus placebo (pbo) in patients with advanced/metastatic renal cell carcinoma (mRCC): Updated safety results. *J Clin Oncol.* 2011;29(7 Suppl1):313.[86]

5. Motzer RJ, Nosov D, Eisen T, et al. Tivozanib versus sorafenib as initial targeted therapy for patients with metastatic renal cell carcinoma: results from a phase III trial. *J Clin Oncol* 2013;31(30):3791-9.[63]

6. Remak E, Charbonneau C, Negrier S, Kim ST, Motzer RJ. Economic evaluation of sunitinib malate for the first-line treatment of metastatic renal cell carcinoma. *J Clin Oncol.* 2008;26(24):3995-4000.[87]

7. Parasuraman S, Hudes G, Levy D, Strahs A, Moore L, DeMarinis R. Comparison of quality-adjusted survival in patients with advanced renal cell carcinoma receiving first-line treatment with temsirolimus (TEMSR) or interferon- (IFN) or the combination of IFN+TEMSR. *J Clin Oncol.* 2008;25(18Suppl):5049.[88]

8. Motzer RJ, Hutson TE, Tomczak P, et al. Sunitinib versus Interferon Alfa in Metastatic Renal-Cell Carcinoma. *New England Journal of Medicine.* 2007;356(2):115-124.[28]

9. National Institute for Health and Care Excellence. NICE Process [PMG9]. Guide to the methods of technology appraisal 2013. National Institute for Health and Care Excellence 2013 April 4 . Available from: http://nice.org.uk/process/pmg9.[68]

Intervention technology and comparators

Table 32 Marketing authorisations

Treatment	Indication
Cabozantinib	It is anticipated that cabozantinib will be indicated for the treatment of advanced RCC in treatment-naïve adults with intermediate or poor risk per IMDC criteria.
Sunitinib	Sunitinib is indicated for the treatment of advanced and metastatic RCC.[89]
Pazopanib	Pazopanib is indicated for the first-line treatment of advanced RCC and for patients who have received prior cytokine therapy for advanced disease.[90]

Abbreviations: IMDC, International Metastatic Renal Cell Carcinoma Database Consortium; RCC, renal cell carcinoma;

Cabozantinib and sunitinib are implemented in the model as per the dosing schedule observed in the CABOSUN study, and as described in Table 9. Cabozantinib was given orally once a day at 60 mg. Sunitinib was administered orally once per day at a dose of 50 mg for 4 weeks, followed by a 2-week break. Treatment modifications,

including interruptions and dose reductions, were used in the CABOSUN study to manage AEs. The dose of cabozantinib could be reduced to 40 mg and then 20 mg, and the dose of sunitinib could be reduced to 37.5 mg and then 20 mg. Dose reductions occurred for 36 (46%) patients in the cabozantinib group and 25 (35%) patients in the sunitinib group (see Section B.2.10). Treatment was continued until disease progression, intolerance to therapy, or withdrawal of consent for treatment. Crossover between treatment arms was not prescribed by the protocol of the CABOSUN clinical trial.[2] The standard daily dose for pazopanib was 800 mg/day (400 mg twice daily) in the COMPARZ clinical trial.[27] Total cost per patient for cabozantinib, sunitinib and pazopanib was adjusted for dose intensity.

Treatment continuation rule

Time to treatment discontinuation (TTD) data from the CABOSUN study were used in the model to inform the comparison of cabozantinib and sunitinib. Parametric survival curves estimated from the CABOSUN patient-level data show the duration of treatment to be different to PFS across both study arms. TTD data for pazopanib was not identified through the systematic literature review and certain assumptions were made to derive the pazopanib TTD (See Section B.3.3, page 93).

B.3.3 Clinical parameters and variables

Incorporation of clinical data in the model/overall survival

The pivotal study to inform the comparison between cabozantinib and sunitinib was the CABOSUN study, described in detail in Section B.2. The OS and PFS data from the CABOSUN study were used to calculate the proportion of patients in each treatment arm at any time point after starting treatment. The proportion of patients in the post-progression health state at any given time was calculated as the difference between OS and PFS. Because there are no head-to-head trials comparing cabozantinib with pazopanib, an ITC was performed (see Section B 2.9). Table 33 summarises the key input options for efficacy data in the model.

Table 33 Summary of key efficacy model input parameters

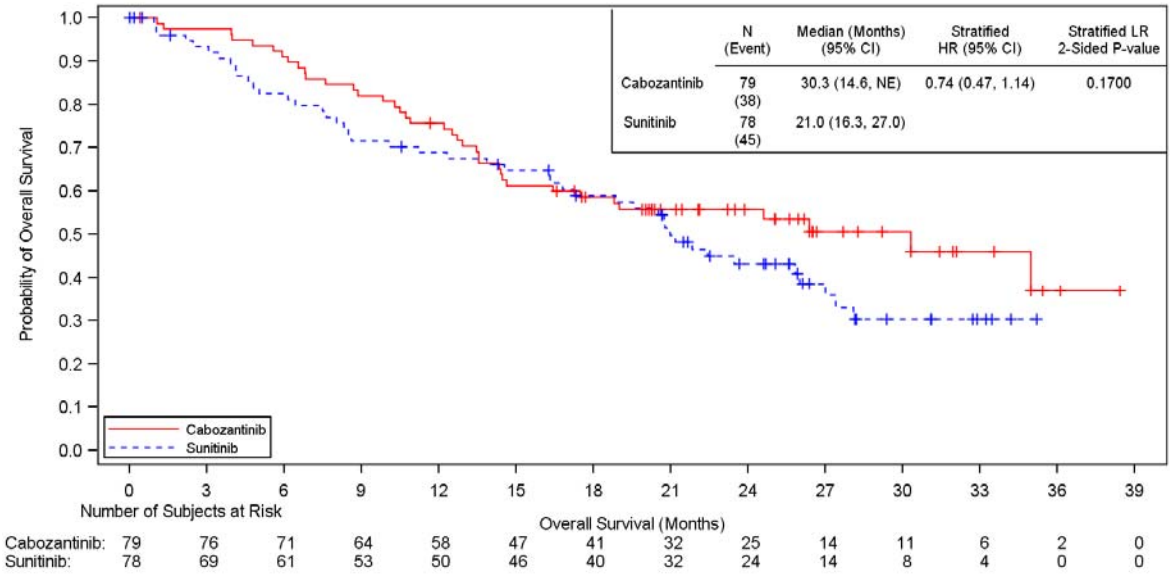
	Model input	
	Cabozantinib vs. sunitinib	**Cabozantinib vs. sunitinib** **Cabozantinib vs. pazopanib**
Efficacy	Patient-level data in CABOSUN study	Regenerated data from the CABOSUN and COMPARZ studies and adjusted efficacy curves of pazopanib to CABOSUN study.
Distributions fitted efficacy data (PFS, OS)	Exponential Weibull Log-logistic Log-normal Gompertz Generalized gamma	Method from Ouwens et al.[91]: Exponential Weibull Log-logistic Log-normal Gompertz Method from Jansen et al.[92] (fractional polynomial method): 1st order model, P= 0 1st order model, P= 0.5 1st order model, P= 1 1st order model, P= -0.5 1st order model, P= -1 2nd order model, P1=0.5, P2=0 2nd order model, P1=-1, P2=0 2nd order model, P1=-1, P2=0.5 2nd order model, P1=-1, P2=1 2nd order model, P1=-1, P2=-1
Distributions fitted efficacy data (TTD)	Exponential Weibull Log-logistic Log-normal Gompertz Generalized gamma	No TTD data for pazopanib was identified through the systematic literature review. Thus, the ITC cannot be applied for TTD data.
Best fitted distributions for PFS, OS and TTD	PFS: log-normal OS: exponential TTD: log-normal	PFS: fractional polynomial 2nd order model, P1=-1 and P2=-1 OS: fractional polynomial 2nd order model, P1=-1 and P2=-1 TTD: log-normal

Abbreviations: BSC, best supportive care; PFS, progression-free survival; OS, overall survival, TTD, time to treatment discontinuation

Overall survival

Patient-level data from the CABOSUN study were used to estimate OS in the cabozantinib and sunitinib arms of the model. Figure 13 shows the Kaplan-Meier OS data from the CABOSUN study, including the number of patients who were at risk or were censored over time. Parametric survival models (exponential, Weibull, Gompertz, log-logistic, log-normal and generalised gamma) were fitted to the patient-level data from the CABOSUN study. Treatment was also tested as a covariate in these parametric models. To select the best survival model fit, the algorithm (SMEEP) as described in the NICE DSU Technical Support Document 14[68] was followed. This included the use of Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC) statistics (Table 34 and Table 35); visual inspection of the curves. The plausibility of different extrapolations was also assessed by visual inspection by oncologists currently practising within the NHS in England. The most appropriate model was selected based on a combination of all these factors.

Figure 13 Kaplan-Meier plot of overall survival (ITT)



Source: Clinical Study Report (OS cut-off January 13th 2017)[44]

The results from the ITC were used to inform OS estimates for the pazopanib comparison (see Section B 2.9). Exponential, Weibull, Gompertz, log-logistic, log-normal and generalised gamma distributions were planned for inclusion in the ITC. However, generalised gamma distribution was not implemented because of the unavailability of the incomplete gamma function in Winbugs. In addition to the ITC

including curves as described by Ouwens, the fractional polynomial ITC was also applied to generate the curve parameters. Five first order and five second order models were fitted to OS data, and the statistical fit is shown in Table 23 and Table 24.

The assumptions of proportional hazards (PH) and accelerated failure time were tested (see ITC methods, Appendix D), to assess whether survival analysis stratified by treatment group was appropriate. Patient-level data generated from publications of included studies suggest that the PH assumption holds for COMPARZ OS data, but not for CABOSUN OS data. It was decided that separate fits were to be used, as the choice of ITC method does not require the PH assumption to hold.

The separately fitted log-normal model was the best fit for cabozantinib in the CABOSUN patient-level data, and the Gompertz model was the best fit to sunitinib, closely followed by exponential. However, oncologists consulted considered that the log-normal was not a reasonable assumption for long-term OS extrapolation as it had a high survival rate at Year 10 (10% for cabozantinib and 8% for sunitinib); instead, the exponential distribution had a more reasonable survival rate at Year 10. Log-logistic and gamma distributions were not used for the same reason as log-normal, i.e. oncologists considered the long-term survival to be too optimistic. Considering that fitting cabozantinib and sunitinib into different curves results in differently shaped distributions, which is not recommended, the exponential distribution was chosen as the base case for OS efficacy data for the comparison between cabozantinib and sunitinib. Also, the scenario analyses were tested using Weibull and Gompertz distributions.

Based on the statistical fit of DIC in the ITC, the fractional polynomial (FP) 2nd order model (P1=-1, P2=-1) distribution provided the best fit for the CABOSUN and COMPARZ re-generated data. Therefore, FP 2nd order (P1=-1, P2=-1) models were used for the comparisons between cabozantinib and sunitinib or pazopanib. In addition, only two FP 2nd order models (P1=-0.5, P2=0) and (P1=-1, P2=0) were tested in the scenario analyses, as the other FP 2nd order models have an unreasonably flat tail. The FP 1st order models have lower DIC than the FP 2nd models, and therefore were not included in the scenario analyses. The exponential,

Weibull and Gompertz distributions of parametric survival curves with RE were included in the scenario analyses. The statistical fits for CABOSUN patient-level analyses are shown in Table 34 and Table 35, and the statistical fits for ITC analyses are shown in Table 23 and Table 24.

Table 34 AIC, AICC and BIC statistics for independently fitted OS data from the CABOSUN study – cabozantinib

Model	AIC	Model	AICC	Model	BIC
Log-normal	353.53	Log-normal	353.69	Log-normal	358.27
Log-logistic	353.97	Log-logistic	354.13	Exponential	358.32
Gamma	355.53	Gamma	355.85	Log-logistic	358.71
Exponential	355.95	Exponential	356.00	Weibull	360.74
Weibull	356.00	Weibull	356.16	Gompertz	362.48
Gompertz	357.74	Gompertz	357.90	Gamma	362.64

Abbreviations: AIC, Aikaike's Information Criterion; AICC, Akaike's Information Criterion (corrected); BIC, Bayesian Information Criterion

Table 35 AIC, AICC and BIC statistics for independently fitted OS data from the CABOSUN study – sunitinib

Model	AIC	Model	AICC	Model	BIC
Gompertz	390.46	Exponential	396.30	Gompertz	394.47
Exponential	396.24	Weibull	397.31	Exponential	398.60
Weibull	397.15	Log-logistic	398.44	Weibull	401.86
Log-logistic	398.28	Log-normal	398.48	Log-logistic	402.99
Log-normal	398.32	Gompertz	398.48	Log-normal	403.04
Gamma	399.10	Gamma	399.42	Gamma	406.17

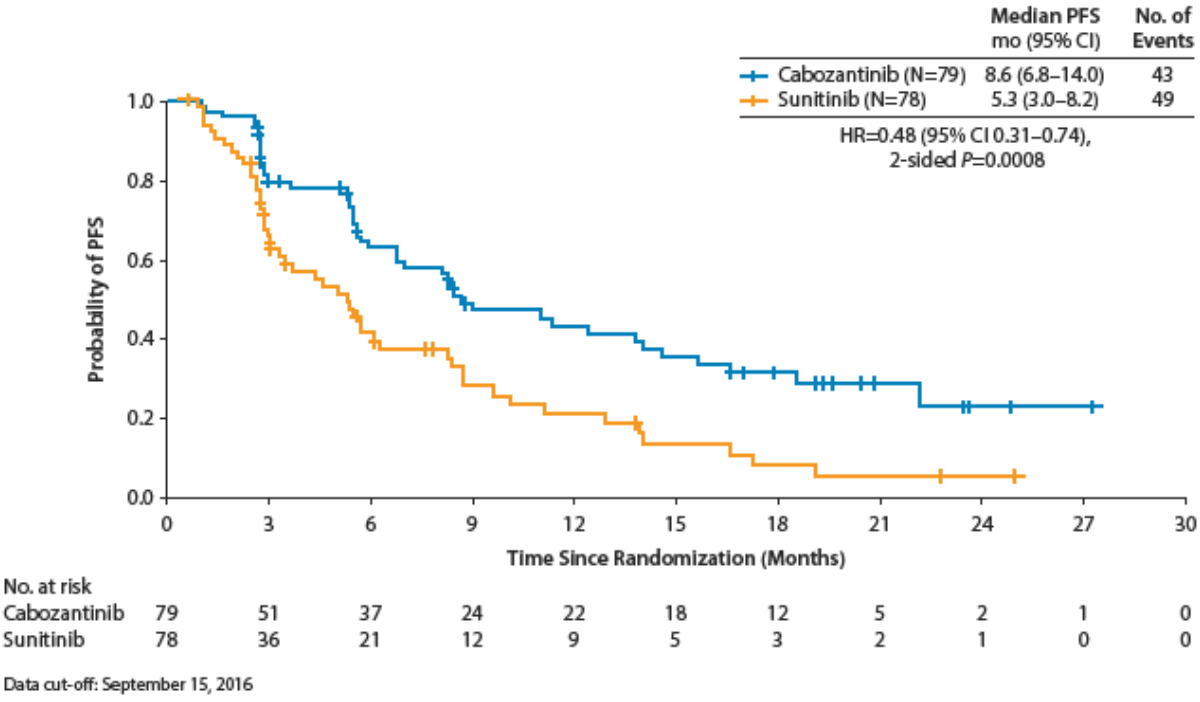
Abbreviations: AIC, Aikaike's Information Criterion; AICC, Akaike's Information Criterion (corrected); BIC, Bayesian Information Criterion

Progression-free survival

Similarly to the OS endpoint, patient-level data from the CABOSUN study were used to inform PFS in the cabozantinib and sunitinib arms of the model. Figure 14 shows the Kaplan-Meier PFS data for CABOSUN patients, including the number of patients who were at risk or were censored over time. Parametric models (exponential, Weibull, Gompertz, log-logistic, log-normal and generalised Gamma) were fitted to the patient-level data from the CABOSUN study. Treatment was also tested as a covariate in these parametric models. To select the best survival model, the

algorithm (SMEEP) as described in the NICE DSU Technical Support Document 14 was followed.[93] The AIC, AICC and BIC statistics are shown in Table 36 and Table 37. Comparing the AIC, AICC and BIC statistics, the log-normal model provided the best fit to the cabozantinib CABOSUN data and the sunitinib data (see Table 36 and Table 37). As with OS, the plausibility of different extrapolations was assessed by visual inspection by oncologists currently practising within the NHS in England. The most appropriate model was identified based on a combination of statistics and visual inspection. The ITC was used to inform PFS estimates of sunitinib and pazopanib comparisons (see Section 4.10).

Figure 14 Kaplan-Meier plot of PFS (ITT)



Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; Mo, months; PFS, progression-free survival; Source: Clinical Study Report[44], Choueiri et al. 2017[47]

Exponential, Weibull, Gompertz, log-logistic, log-normal and generalised gamma were planned to be included in the ITC. However, the generalised gamma distribution was not implemented because of the unavailability of the incomplete gamma function in Winbugs. In addition to the parametric curve ITC, the fractional polynomial ITC was also applied to generate the curve parameters. Five first order

and five second order models were fitted to PFS data, and the statistical fit is shown in Table 23 and Table 24.

The assumptions of proportional hazards and accelerated failure time were tested, to assess whether survival analysis stratified by treatment group was appropriate (ITC methods, Appendix D). Patient-level data generated from publications of included studies suggested that PH assumption holds for COMPARZ PFS data, but not for CABOSUN PFS data. Therefore, separate fits were used, as the choice of ITC method does not require the PH assumption to hold.

Separately fitted log-normal distribution provided the best fit to the cabozantinib and sunitinib PFS data, followed by gamma and log-logistic distribution. Considering that fitting cabozantinib and sunitinib into different curves results in differently shaped distributions, which is not recommended by the NICE DSU Technical Support Document 14,[93] the log-normal distribution was chosen in the base case analysis for PFS efficacy data for the comparison between cabozantinib and sunitinib. In addition, the exponential, Weibull and Gompertz distributions were included into the scenario analyses.

The FP 2nd order model (P1=-1, P2=-1) provided the best fits to the CABOSUN patient-level data and COMPARZ re-generated data compared to other distributions. The FP 2nd order model (P1=-1, P2=-1) was used in the base case for the comparisons between cabozantinib and sunitinib or pazopanib. As the other FP 2nd order models have an high PFS rate at Year 5 , no other FP 2nd order models were included in scenario analyses. The exponential, Weibull and Gompertz distributions of parametric survival curves with random effect were included into the scenario analyses, whereas the statistical fits for ITC analyses are shown in Table 23 and Table 24.

Table 36 AIC, AICC and BIC statistics for independently fitted PFS data from CABOSUN study - cabozantinib

Model	AIC	Model	AICC	Model	BIC
Log-normal	317.16	Log-normal	317.34	Log-normal	321.69
Gamma	317.85	Gamma	318.20	Log-logistic	323.55
Log-logistic	319.02	Log-logistic	319.20	Gamma	324.63
Exponential	323.31	Exponential	323.37	Exponential	325.57
Weibull	324.01	Weibull	324.18	Weibull	328.53
Gompertz	325.28	Gompertz	325.46	Gompertz	329.80

Abbreviations: AIC, Aikaike's Information Criterion; AICC, Akaike's Information Criterion (corrected); BIC, Bayesian Information Criterion

Table 37 AIC, AICC and BIC statistics for independently fitted PFS data from CABOSUN study - sunitinib

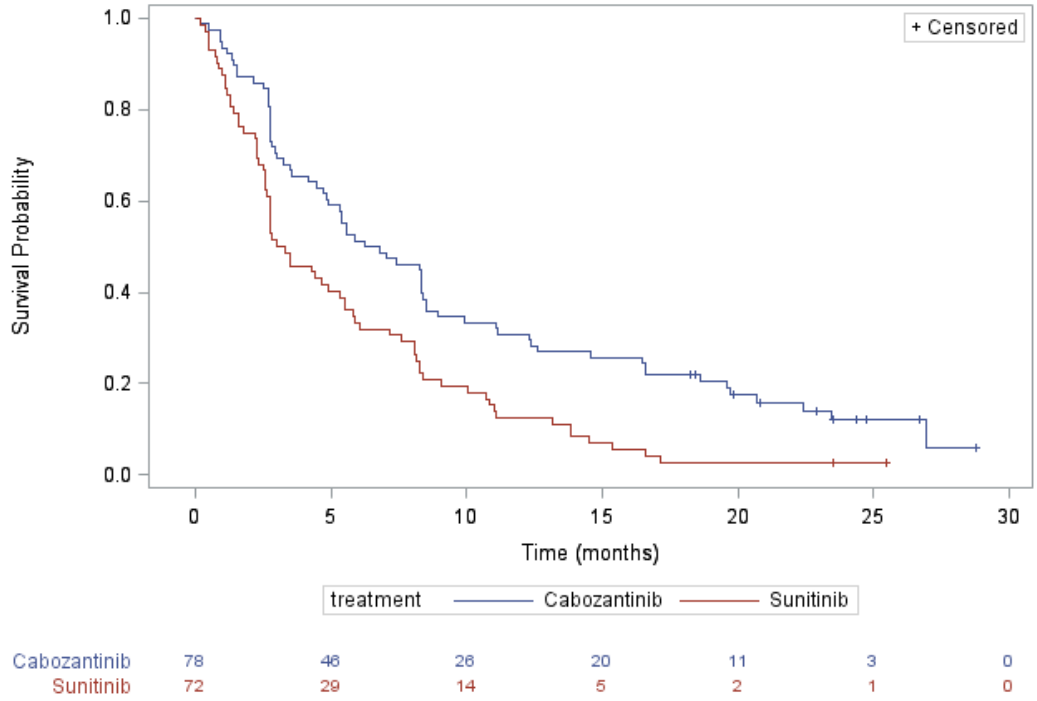
Model	AIC	Model	AICC	Model	BIC
Log-normal	291.52	Log-normal	291.71	Log-normal	295.83
Gamma	292.37	Gamma	292.77	Log-logistic	297.98
Log-logistic	293.66	Log-logistic	293.86	Gamma	298.84
Weibull	300.47	Weibull	300.66	Exponential	303.20
Exponential	301.04	Exponential	301.11	Weibull	304.79
Gompertz	303.00	Gompertz	303.20	Gompertz	307.32

Abbreviations: AIC, Aikaike's Information Criterion; AICC, Akaike's Information Criterion (corrected); BIC, Bayesian Information Criterion

Time to discontinuation (TTD)

In the economic model, TTD determined the proportion of patients on treatment at each point in time. For cabozantinib and sunitinib, treatment duration was based on TTD data from the CABOSUN trial. For pazopanib, no TTD data were obtained from publications. The best fit was identified as described for OS and PFS.

Figure 15 Kaplan-Meier plot of TTD



According to the AIC, AICC or BIC statistics, the log-logistic distribution provided the best fit to the cabozantinib CABOSUN data, followed by the log-normal distribution, while the log-normal provided the best fit for the sunitinib CABOSUN data (see Table 38 and Table 39). Considering that fitting cabozantinib and sunitinib into different curves results in differently shaped distributions, which is not recommended within the NICE DSU Technical Support Document 14[93], the log-normal distribution was used in the base case for cabozantinib and sunitinib.

Table 38 AIC, AICC and BIC statistics for independently fitted TTD data from CABOSUN study - cabozantinib

Model	AIC	Model	AICC	Model	BIC
Log-logistic	459.07	Log-logistic	459.23	Exponential	463.54
Log-normal	459.31	Log-normal	459.47	Log-logistic	463.79
Gamma	460.87	Gamma	461.20	Log-normal	464.02
Exponential	461.18	Exponential	461.24	Gompertz	467.26
Gompertz	462.54	Weibull	463.33	Weibull	467.89
Weibull	463.17	Gompertz	463.33	Gamma	467.94

Abbreviations: AIC, Aikaike's Information Criterion; AICC, Akaike's Information Criterion (corrected); BIC, Bayesian Information Criterion

Table 39 AIC, AICC and BIC statistics for independently fitted TTD data from CABOSUN study - sunitinib

Model	AIC	Model	AICC	Model	BIC
Log-normal	387.21	Log-normal	387.38	Exponential	390.94
Gamma	388.43	Exponential	388.72	Log-normal	391.76
Log-logistic	388.56	Log-logistic	388.73	Log-logistic	393.11
Exponential	388.66	Gamma	388.79	Weibull	394.96
Weibull	390.41	Weibull	390.58	Gompertz	394.99
Gompertz	390.43	Gompertz	390.61	Gamma	395.26

Abbreviations: AIC, Aikaike's Information Criterion; AICC, Akaike's Information Criterion (corrected); BIC, Bayesian Information Criterion

For pazopanib, no TTD data were found from the literature search. It was assumed that pazopanib TTD is equal to sunitinib TTD. This assumption was based on the COMPARZ study, where mean treatment duration for both sunitinib and pazopanib was 11.5 months.[94]

Changes to transition probabilities over time/expert opinion

The relative effectiveness between treatments is based on head-to-head comparison for cabozantinib and sunitinib, and on the ITC for pazopanib. The curves are used beyond the clinical trial duration, until the end of the model time horizon. This is associated with uncertainty. Different distribution types were tested in scenario analyses.

B.3.4 Measurement and valuation of health effects

The symptoms associated with advanced RCC and the generally poor prognosis can significantly affect all domains of patients' HRQL including physical function and psychosocial wellbeing.[7,8] Symptoms of advanced RCC include fatigue, weight loss, anaemia, and paraneoplastic syndromes involving hypertension, fever, cachexia (wasting), neuromyopathy and amyloidosis.[15] Additional symptoms related to the metastatic spread of the disease include bone pain, skeletal-related events and hypercalcaemia; lung symptoms such as airway obstruction; and venous thromboembolism.[14,16]

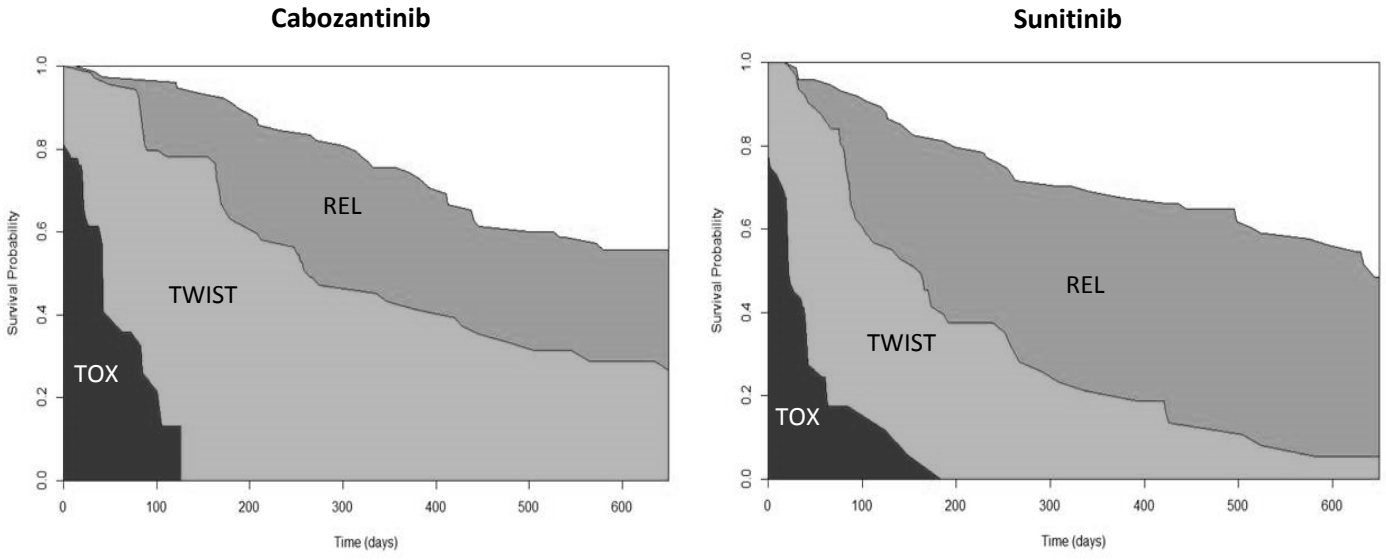
Health-related quality-of-life data from clinical trials

The CABOSUN trial did not collect EQ-5D or any other generic preference-based measure to estimate utilities. Other quality of life data to enable mapping to a generic

measure were also not collected. Hence, no utility estimates from CABOSUN are included in this model.

Although not used in the model, a post hoc analysis using a quality-adjusted time without symptoms of disease or toxicity of treatment (Q-TWiST) methodology was conducted. Each patient's overall survival was partitioned into three mutually exclusive health states: grade 3 or 4 toxicity (TOX), time without symptoms of disease or grade 3/4 toxicity (TWiST), and time after progression or relapse (REL). Time spent in each state was weighted by a health-state utility associated with that state and summed to calculate the Q-TWiST. A threshold utility analysis was used, applying utilities across the range of 0 (similar to death) to 1 (perfect health). The analysis period was 650 days (median survival follow-up period). The mean time spent with TWiST was 121 days (95% confidence interval [CI], 43–199) higher for cabozantinib compared with sunitinib. The mean time spent with TOX was 8 days higher for cabozantinib, while the mean time spent with REL was 104 days shorter for cabozantinib. The results are shown in Figure 16. In the threshold utility analysis, the difference in Q-TWiST ranged from 129 days (utility TOX=1, REL=0) to 17 days (TOX=0, REL=1), in favour of cabozantinib across all the utility combinations. Patients treated with cabozantinib were associated with longer Q-TWiST compared to those treated with sunitinib, primarily due to longer time without symptoms of disease or grade 3/4 toxicities.

Figure 16 Q-TWiST partitioned survival curves



Health-related quality-of-life studies

Systematic searches for relevant HRQL data are described in Appendix H

Systematic literature review of utility studies

A systematic literature review was conducted to compare the utility values derived from published articles. This review was an update of a previous systematic literature review carried out in 2016.[38] The previous review, comprising studies evaluating first- and second-line treatment of advanced and metastatic RCC, found 57 studies related to both HRQL and utility. An update was performed aimed at searching for utilities (i.e. not HRQL studies). The previous search is described in detail in NICE TA463.[38] The update of the systematic review is reported in detail in Appendix H.

Key differences in utility values

Combining the publications identified from the systematic literature review in July 2016 and the publications identified from the new systematic literature review in August 2017, a total of 22 utility publications were identified which report EQ-5D utility values for first-line treatment of advanced and metastatic RCC.

The studies relevant for the UK setting are presented in Table 40. Additionally, the complete list of studies identified by Systematic literature review is reported in Appendix H, Table 30. Most studies either reported the treatment-specific utility values, or reported the utility values designed for a cost-effectiveness study in a country other than the UK. In the previous NICE appraisals for advanced RCC,[82 35] treatment-specific utility values were not considered reasonable assumptions, and the health-state-specific utility values were preferred instead. Also, the recent DSU report[95] stated that utility values in the UK were preferred to those in other countries.

Swinburn et al. conducted a study on elicitation of health state utilities in metastatic RCC in 2010. In this study, health state descriptions were developed based on a literature review and in-depth interviews with clinical experts. The states included description of stable disease, progressive disease, and toxicities that may be experienced by patients receiving treatments for their advanced RCC. Opinions on health-state-related utility values from the general public were elicited using the time-trade off (TTO) method.

Additionally, previous NICE submissions for key comparators (tivozanib, pazopanib, and sunitinib) were checked. The utility values from key HTA submissions are summarised in Table 41, Source: Tivozanib NICE Technology Appraisal17

Table 42, and Table 43. The tivozanib NICE submission used the utility values from Motzer 2013.[63]

Table 40 Reported utility values in studies evaluating first line RCC therapy

Study	Health state	Population	Interventions	Results	Comments
Studies identified in the SLR conducted in August 2017					
Amdahl, 2017[77]	PFS and mixed health states	UK	Pazopanib Sunitinib	Regression model-derived EQ-5D: Pazopanib pre-progression: 0.817, post-progression: 0.785; Sunitinib pre-progression: 0.797, post-progression: 0.746;	Utilities of health states are treatment specific.
Studies identified in the SLR conducted in July 2016					
Castellano, 2009[96]	Baseline and over the first six cycles; PFS	France, Germany, Italy, Spain, UK, Poland	IFNα Sunitinib	EQ-5D: *Baseline, mean (SD):* Sunitinib: 0.72 (0.24); IFNα: 0.74 (0.25); *First six cycles, mean:* Sunitinib: 0.723; IFNα: 0.674;	Utilities of health states are treatment specific.
Lai, 2016[97]	PFS	France, Italy, UK, Germany, Finland	Pazopanib Sunitinib	EQ-5D: Baseline mean (SD): 0.77 (± 0.24);	This study only contains the utility for baseline, but not the utility for health states.
Swinburn, 2010[20]	PFS, PD	UK	Various treatment for mRCC	EQ-5D: Stable disease no AE: 0.795 (0.0176); Disease progression: 0.355 (0.0288);	Relevant utilities

Abbreviations: NR, Not reported; BSC, Best supportive care; PFS, Progression-free state; PD, Progression of disease; IFNα, Interferon-alfa; EQ-5D, Euroqol descriptive system questionnaire; SD, standard deviation; SE, standard error; SLR, systematic literature review; mRCC, metastatic renal cell carcinoma

Table 41 GID-TA10123 – tivozanib

State	Utility value (SE)	Comments
Progression-free state	0.726 (0.011)	TIVO-1, EQ-5D-3L The utility values are for the full TIVO-1 study population (i.e. not treatment naïve population). 30% of patients had been pre-treated with prior systemtic therapy. The trial also included 27% and 34% favourable risk patients in the tivozanib and sorafenib arms, respectively.
Progressed disease state	0.649 (0.019)	TIVO-1, EQ-5D-3L The utility values are for the full TIVO-1 study population (i.e. not treatment naïve population). 30% of patients had been pre-treated with prior systemtic therapy. The trial also included 27% and 34% favourable risk patients in the tivozanib and sorafenib arms, respectively.
Abbreviations: SE, standard error.		

Source: Tivozanib NICE Technology Appraisal[17]

Table 42 TA215 - pazopanib

State	Utility value (SE)	Comments
Progression-free state	0.70 (0.01)	Based on the mean EQ-5D utility value among all patients without AEs in the VEG105192 trial
Progressed disease state	0.59 (NA)	Progression was assumed to be associated with a decrement in utility of 15%
Abbreviations: SE, standard error; NA, not applicable.		

Source: Pazopanib NICE Technology Appraisal[35]

Table 43 TA169 - sunitinib

State	Utility value (SE)	Comments
Progression-free	0.780 (NA)	Average EQ5D utility value
Progressed	0.705 (NA)	Average EQ5D utility value
Abbreviations: SE, standard error; NA, not available.		

Source: Sunitinib NICE Technology Appraisal[19]

Adverse reactions

Adverse reactions are considered to be an important driver of quality of life in advanced and metastatic RCC patients. The literature review identified a range of disutility values, see Table 44.

Table 44 Summary of available disutility values

Source	Utility / Disutility (SE)	Details
Amdahl, 2016[98]	-0.2044 (0.0682)	Disutility value for all AEs for grades 3–4 derived from the COMPARZ clinical trial[27]
Goebell, 2014[99]	0.76 (0.2300)	Fatigue
METEOR study[100]	-0.055 (0.0068)	Grade 3 or 4 AEs: METEOR patient-level analyses[38]
Swinburn, 2010[20]	0.676 (NA)	Stable with anaemia grade III
	0.534 (NA)	Stable with diarrhoea grade III
	0.591 (NA)	Stable with fatigue grade III
	0.540 (NA)	Stable with nausea grade III
	0.642 (NA)	Stable with hypertension grade III
	0.469 (NA)	Stable with PPES grade III
TIVO-1 study[17]	0.61 (0.0200)	Anaemia
	0.60 (0.0260)	Asthenia/fatigue
	0.68 (0.0060)	PPES
	0.66 (0.0070)	Hypertension
	0.71 (0.0100)	Diarrhoea
Abbreviations: AE, Adverse event; PPES, Palmar-plantar erythrodysesthesia; NA, not available.		

To capture the effect of the AEs on HRQL, assumptions about the durations of AEs were required. In the current analyses it was assumed that duration of an AE was 4 weeks. In the absence of trial-based information, this assumption was based on the analysis of the study data from the cabozantinib METEOR clinical trial. This approach was used as it can be extended to the comparator (pazopanib) where no direct evidence exists. Table 45 shows the duration of disutility and the QALY decrements associated with each AE.

Table 45 Inputs of adverse events in the cost-effectiveness model

State	Duration of adverse events	Number of episodes experienced per patient	QALY decrement	Source
TEAE, grade 3/4	4 weeks – assumption	1.429	-0.2044	Amdahl, 2016[98]
Abbreviations: QALY, quality adjusted life year; TEAE, treatment emergent adverse event				

Health-related quality-of-life data used in the cost-effectiveness analysis

Given the difficulties in combining utility estimates from different sources, including differences in trial populations and/or elicitation methods, and the lack of patient-reported data from CABOSUN, the base case analysis uses utility values derived from the TIVO-1 study as reported in the tivozanib NICE technology appraisal.[17] The values used are 0.726 for pre-progression, and 0.649 for post-progression. Clinicians suggested that utility values for sunitinib reported in the sunitinib NICE technology appraisal[19] should be used, as it is the direct comparison of cabozantinib in the CABOSUN trial. However, in the cost-effectiveness analysis the utility data from TIVO-1 study is used, as it is the latest appraised data by NICE. Scenario analyses are provided using the utilities from the pazopanib NICE appraisal,[18] sunitinib NICE appraisal[19] and Swinburn 2010.[20]

Changes to HRQL over course of disease

Although it is possible that the patient's utility might vary during progression, a single mean value is used in this analysis to represent the whole health state. However, the utility values in the cost-effectiveness model change between health states (i.e. separate utility values for the progression-free and progressed disease health states).

Baseline utility values

Baseline quality of life was not directly assumed in the economic evaluation as all patients start from the progression-free health state, and they either remain in this state or transition to the progressed disease state.

Adjustment of utility values

As age increases, the utility values decrease in the general population. Thus, age-adjusted utility values were included in the cost-effectiveness analysis to take this effect into consideration.[101]

Excluded health effects

No other health effects found in the literature or identified in clinical trials were intentionally excluded from the cost-effectiveness analysis.

Summary of utility values in cost-effectiveness analysis

The utility values are shown in Table 46 for the base case analysis and Table 47 for the scenario analysis.

Table 46 Summary of utility values for cost-effectiveness analysis – base case

State	Utility value: mean (SE)	95% CI	Reference in submission (section and page number)	Justification
Progression-free	0.726 (0.011)	0.705; 0.748	Section B.3.4, Table 41, page 100	Latest utility values accepted by the NICE Appraisal Committee (tivozanib appraisal, TA10123).[17]
Post-progression	0.649 (0.019)	0.612; 0.686		
TEAEs Grade 3/4	-0.2044 (0.0682)	-0.0707; -0.3381	Section B.3.4, Table 45, page 102	COMPARZ study data, mean decrement due to Grade 3/4 adverse event. This is also the lowest value for AE related disutilities and is a conservative assumption.
Abbreviations: CI, confidence interval; SE, standard error; TEAE, treatment emergent adverse events.				

Table 47 Summary of utility values for cost-effectiveness analysis – scenario analysis

State	Utility value: mean (SE)	95% CI	Reference in submission (section and page number)	Justification
Progression-free survival	0.70 (0.01)	0.680; 0.720	Section B.3.4 Table 42, page 100	In absence of the health state specific utility values from CABOSUN study, the previous utility values that were accepted in the pazopanib NICE appraisal were applied.[18]
Post-progression survival	0.59 (0.059)*	0.474; 0.706*		

Progression-free survival	0.78 (0.078)*	0.627; 0.933*	Section B.3.4 Table 43 page 100	In absence of the health state specific utility values from CABOSUN study, the previous utility values that were accepted in the sunitinib NICE appraisal were applied.[19]
Post-progression survival	0.705 (0.071)*	0.567; 0.843*		
Progression-free survival	0.795 (0.0176)	0.761; 0.830	Section B.3.4 Table 40, page 99	In absence of the health state specific utility values from CABOSUN study, the utility values of UK population from Swinburn 2010 were applied.[20]
Post-progression survival	0.355 (0.0288)	0.299; 0.412		
Treatment-emergent adverse events, grade 3 & 4	-0.055 (0.0068)	-0.0418; -0.0685	Section B.3.4 Table 44, page 101	In the absence of AE-specific data from CABOSUN study, the utility value obtained from the METEOR clinical study which assessed cabozantinib and everolimus in the second line setting for RCC, was tested in the scenario analysis.[69]
Abbreviations: CI, confidence interval; SE, standard error; NA, not available. *SE or 95% CI were not available in the literature and 10% of the mean values were assumed to be the SE.				

B.3.5 Cost and healthcare resource use identification, measurement and valuation

Appendix I describes how relevant cost and healthcare resource data were identified.

Intervention and comparators' costs and resource use

Drug and treatment costs

Table 48 shows the treatment acquisition costs for cabozantinib and its comparators, pazopanib and sunitinib, based on the NHS list prices. Costs are displayed per week, representing the one week model cycle. Table 48 shows the cost of treatment. All drug costs are based on NHS list prices. As cabozantinib, pazopanib and sunitinib are all oral drugs, no cost of administration was applied.

For each treatment, drug cost per cycle is determined based on dosage after accounting for relative dose intensity (Table 48). For cabozantinib, a flat price for each dose (20, 40, and 60 mg) was applied, consequently any dose reduction does not impact on the cost. However, some patients have dose interruptions, which means that

a standard pack of 30 cabozantinib tablets can last for more than 30 days. In order to take this into consideration in the economic model, the percentage of days with dose interruption was applied. The percentage of days not interrupted is 94.3% (mean), an estimate based on CABOSUN trial data. The dose intensities of sunitinib and pazopanib are 87%[44] and 86%,[85] respectively.

The NICE guidance for pazopanib (TA463) states that the manufacturer provides pazopanib with a 12.5% discount on the list price as agreed in the patient access scheme (PAS).[35] Also, the manufacturer of sunitinib agreed a PAS with the Department of Health, in which the first treatment cycle of sunitinib is free to the NHS.[19] These PASs were applied in the base case, together with the PAS which is in place for cabozantinib for the second-line treatment of RCC (TA 463) and which will also apply to the first-line setting.[38]

Table 48 Drug formulation, dose and total cost per week for 1st line treatments

Drug	Formulation (mg)	Cost per pack, £	Tablets per admin	Tablets per pack	Dose, mg	Weekly frequency	Relative dose intensity, % (SE)	Total cost per model cycle including discount, £
Cabozantinib	20/40/60	5143	1	30	60/20/40	7	94.3 (1.5)*	[redacted]
Sunitinib	50	3138	1	28	50	4.7**	87.4 (6.3)2	457.22
Pazopanib	400	1121	2	30	800	7	86.0 (8.6)3	393.66

Sources:

1 British National Formulary (BNF), NHS indicative price.[102]

2 CABOSUN CSR, Table 37, sunitinib mean relative dose intensity.[44]

3 NICE Single technology appraisal. Pazopanib for the first-line treatment of patients with advanced renal cell carcinoma (RCC).[18]

Notes:

* Dose interruptions due to adverse events: 94.3% (mean).

** Sunitinib is given in 6 week cycles of 4 weeks of treatment followed by a rest period of 2 weeks.

Health-state unit costs and resource use

The base case and scenario analysis resource use and unit cost estimates attributed to disease management are shown in Table 49 and Table 50. The costs are presented per model cycle (1 week). The health resource utilisation in the base case was estimated by clinicians currently practising in the UK. Resource use assumptions mirror those in the NICE STA for pazopanib[35] and the ongoing tivozanib appraisal.[82]

Prior to therapy initiation, it is assumed that patients require a first consultant appointment with a medical oncologist. Thereafter, follow-up consultant visits and blood tests are assumed to take place on a monthly basis. CT scans are assumed to happen every 3 months. The same resource use is assumed regardless of health state.

In scenario analyses, different disease monitoring assumptions were assumed based on UK clinical opinion. Follow-up consultant visits took place less frequently (every 12 weeks). Additional nurse visits took place (every 4 weeks). The frequency of CT scans and blood tests remained the same.

After the failure of first-line treatment, patients move to second-line treatment. It was assumed that monitoring requirements in post-progression phase will remain the same as in pre-progression phase.

Table 49 Disease management - cost and resource use (base case analysis)

Disease state	Resource	Frequency per cycle	Unit cost (SE), £	Reference
Progression-free	Outpatient consultation (first attendance)	Not applicable	219.0 (21.9)	NHS Reference Costs 2016/2017[103] Currency code WF01B, Service code 370, Medical oncology
	Outpatient consultation (follow up)	0.25	173.0 (17.3)	NHS Reference Costs 2016/2017[103]. Currency code: WF01A, Service code 370, Medical oncology
	CT scan	0.08	115.0 (11.5)	NHS Reference Costs 2016/2017[103]. Currency code: RD25Z Computerised Tomography Scan of three areas, without contrast
	Blood test	0.25	3.0 (0.3)	NHS Reference Costs 2016/2017[103]. Currency code: DAPS05
Progression	Outpatient consultation (follow up)	0.25	173.0 (17.3)	NHS Reference Costs 2016/2017[103]. Currency code: WF01A, Service code 370, Medical oncology
	CT scan	0.08	115.0 (11.5)	NHS Reference Costs 2016/2017[103]. Currency code: RD25Z Computerised Tomography Scan of three areas, without contrast
	Blood test	0.25	3.0 (0.3)	NHS Reference Costs 2016/2017[103]. Currency code: DAPS05
End of life costs	Various	One-off cost*	6207.6 (7.55)	Cost of end of life care*

Abbreviations: CT, computerised tomography. Source: *Theo Georghiou and Martin Bardsley. Exploring the cost of care at the end of life. September 2014. Nuffield Trust.[104]

Table 50 Disease management - cost and resource use (scenario analysis)

Disease state	Resource	Frequency per cycle	Unit cost (SE), £	Reference
Progression-free	Outpatient consultation (first attendance)	Not applicable	219.0 (21.9)	NHS Reference Costs 2016/2017[103] Currency code WF01B, Service code 370, Medical oncology
	Outpatient consultation (follow up)	0.08	173.0 (17.3)	NHS Reference Costs 2016/2017[103]. Currency code: WF01A, Service code 370, Medical oncology
	Nurse visits	0.25	36.0	Cost per hour. Nurse (GP practice), Unit cost of health and social care 2016[105]
	CT scan	0.08	115.0 (11.5)	NHS Reference Costs 2016/2017[103]. Currency code: RD25Z Computerised Tomography Scan of three areas, without contrast
	Blood test	0.25	3.0 (0.3)	NHS Reference Costs 2016/2017[103]. Currency code: DAPS05
Progression	Outpatient consultation (follow up)	0.08	173.0 (17.3)	NHS Reference Costs 2016/2017[103]. Currency code: WF01A, Service code 370, Medical oncology
	Nurse visits	0.25	36.0	Cost per hour. Nurse (GP practice), Unit cost of health and social care 2016[105]
	CT scan	0.08	115.0 (11.5)	NHS Reference Costs 2016/2017[103]. Currency code: RD25Z Computerised Tomography Scan of three areas, without contrast
	Blood test	0.25	3.0 (0.3)	NHS Reference Costs 2016/2017[103]. Currency code: DAPS05
End of life costs	Various	One-off cost*	6207.6 (7.55)	Cost of end of life care*

Abbreviations: CT, computerised tomography. Source: *Theo Georghiou and Martin Bardsley. Exploring the cost of care at the end of life. September 2014. Nuffield Trust.[104]

Adverse reaction unit costs and resource use

The health care costs for the most frequent (≥ 5%) grade 3 and 4 TEAE experienced by cabozantinib and the comparator treatments were included in the cost-effectiveness analysis. The included AEs are described in Section B.2.10. A systematic literature review revealed limited published data on resource use associated with treatment of AEs included in the cost-effectiveness models (see Appendix I). Resource use was estimated based on clinical opinion, published sources and HTA reports from previous NICE RCC appraisals.[35,82]

Inpatient and outpatient costs were obtained from England NHS Payment by Result (PbR) tariffs, and NHS reference costs and drug costs were taken from the British National Formulary (BNF)[102]. The total cost for each AE was obtained by summing the costs of each resource used in managing the AE (i.e. inpatient, day case, outpatients and medication costs) (Table 51 and Table 52). The unit costs identified are presented in Table 53.

Table 51 Unit costs for health resources utilised in management of adverse events

Health resource	Cost, £	Reference[103]
Short stay admission due to diarrhoea	558	HRG FD02E, Inflammatory Bowel Disease without Interventions, with CC Score 5+, NHS reference costs 2016/2017[103]
Vascular ultrasound scan	75	HRG RD47Z, NHS reference costs 2016/2017[103]
Outpatient attendance due to palmar-plantar erythrodysesthesia	173	HRG WF01A: service code 370 Medical oncology, NHS reference costs 2016/2017[103]
Visit to endocrinologist due to hyperglycaemia	146	WF01A, Service code 302, Endocrinology, NHS reference costs 2016/2017[103]
GP visit due to hypertension	36	GP visit-Unit cost per surgery consultation; PSSRU Cost of health and social care 2016[105]
Hospitalisation cost due to lymphocytopenia - Short stay emergency tariff	429	HRG SA35A-SA35E short stay emergency tariff (weighted average), NHS reference costs 2016/2017[103]
Hospitalisation cost due to lymphocytopenia - Day case tariff	330	HRG SA35A-SA35E day case tariff (weighted average), NHS reference costs 206/2017[103]
Hospitalisation cost due to thrombocytopenia - Short stay emergency tariff	522	HRG SA12G-SA12K short stay emergency tariff (weighted average), NHS reference costs 2016/2017[103]
Hospitalisation cost due to Thrombocytopenia - Day case tariff	308	HRG SA12G-SA12K day case tariff(weighted average), NHS reference costs 2016/2017[103]
Outpatient attendance due to dyspnoea	68	WF01A, Service code 342, Programmed Pulmonary Rehabilitation, NHS reference costs 2016/2017[103]
Outpatient visit for pain management	138	WF01A, Service code 191 , Pain management, NHS reference costs 2016/2017[103]

Abbreviations: GP, general practitioner

Table 52 Medication costs for management of the adverse events

Unit	Cost (£)	Package	Reference[103]
Loperamide 2 mg	2.15	30 tbl	BNF (accessed 08.12.2017)
Metformin 500 mg	0.77	28 tbl	BNF (accessed 08.12.2017)
Ramipril 5 mg	0.98	28 tbl	BNF (accessed 08.12.2017)
Bendroflumethiazide 2.5 mg	0.58	28 tbl	BNF (accessed 08.12.2017)
Neupogen 30million units/1 ml solution for injection vials	264.00	5 vials	BNF (accessed 08.12.2017)
Dexamethasone 2 mg/5 ml oral solution	42.30	150 ml	BNF (accessed 08.12.2017)
Deltaparin (Fragmin 18,000 units/0.72 ml)	50.82	5 inj.	BNF (accessed 08.12.2017)
Deltaparin (Fragmin 15,000 units/0.6 ml)	42.34	5 inj.	BNF (accessed 08.12.2017)

Abbreviations: BNF, British National Formulary; Inj, injections; tbl, tablets

Table 53 Management cost for adverse events

Adverse event	Resource Use Assumption	Cost, £
Diarrhoea	Based on Pazopanib NICE STA[1] -Short stay admission[2] and -Loperamide 2 mg[102] q.i.d 30 days	567.0
Dyspnoea	Based on the assumption: One pulmonologist visit[2]	68.0
Embolism	Based on NICE guidance on venous thromboembolic diseases: diagnosis, management and thrombophilia testing[4]: -1 ultrasound of coronary vasels[2] to determine where embolism occurred -Therapy initiation with low molecular weight heparin (LMWH) for 6 months: deltaparin 18000 units o.d. units for first 30 days and continue with deltaparin 15000 units o.d. for further 5 months[4]	1640.0
Fatigue	Based on Tivozanib NICE STA[5] 20% of patients will have additional outpatient attendance[2]	35.0
Hyperglycaemia	Based on assumption: -1 visit to endocrinologists[2] -Initiation of therapy with p.o anti-diabetic medication: metformin 500mg[3] o.d. for one year	156.0
Hypertension	Based on Tivozanib NICE STA[5] -3 GP attendances[6] -Ramipril 5 mg + bendroflumethiazide 2.5 mg[3] o.d. for 1 year	128.0
Lymphocytopenia	Based on the assumption: 20% of short stay emergency tariff (weighted average of SA35A-SA35E: £ 515)[2] and 80% of patients with day case tariff (weighted average of SA35B-SA35E: £ 288)[2]	362.0

Neutropenia	Based on the assumption: - Granulocyte colony-stimulating factors (granulocyte CSF): Filgrastim Filgrastim. 5µg/kg for 14 days[4] (dose is 450 µg o.d. for TM=90kg) Neupogen 30million units/1ml (1µg=100000 units)[3]	1107.0
Pain	Based on the assumption: Monthly outpatient visit for pain management	138.0
Palmar-plantar erythrodysesthesia	Based on Tivozanib NICE STA[5]: -60% of patients will have additional outpatient attendance[2]	104.0
Stomatitis	Based on the assumption: -Local therapy for therapy for pain relief, local anaesthetics or other anti-inflammatory preparations We assume that patient will apply oral solution of dexamethasone 2mg/5ml[3]	42.0
Thrombocytopenia	Based on the assumption: 20% of short stay emergency tariff (weighted average of SA12G-SA12K)[2] and 80% of patients with day case tariff (weighted average of SA12G-SA12K)[2]	351.0
Hyponatremia	Based on the assumption: Regular blood tests (already considered under disease management costs)	0.0
Hypophosphatemia	Based on the assumption: Regular blood tests (already considered under disease management costs)	0.0
Hypotension	Based on the assumption: Monthly outpatients visits[2] -These costs are already covered by disease management costs which comprise monthly outpatient visits	0.0
Increased ALT	Based on the assumption: Regular blood tests (already considered under disease management costs)	0.0
Increased AST	Based on the assumption: Regular blood tests (already considered under disease management costs)	0.0

Cost references:

1. Pazopanib (Votrient®) for the first-line treatment of patients with advanced renal cell carcinoma (RCC) NICE STA[18]
2. National reference costs 2016/2017[103], weighting as per Tivozanib NICE STA ERG assumption. [17]
3. BNF accessed 08.12.2012[102]
https://www.medicinescomplete.com/mc/bnf/64/
4. NICE guidance on Deep vein thrombosis or pulmonary embolism[106]
https://www.nice.org.uk/guidance/cg144/chapter/recommendations#terms-used-in-this-guideline
5. Tivozanib for treating renal cell carcinoma [ID591] NICE STA[17]
https://www.nice.org.uk/guidance/gid-ta10123/documents/committee-papers
6. PSSRU Cost of health and social care 2016[105]
http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2016/

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; od, once daily; qid, four times a day; STA, single technology appraisal

Miscellaneous unit costs and resource use

The model also allows for the inclusion of the end-of-life costs, which occur in the same cycle as death. Cost estimates were taken from a 2014 report on the cost of care at the end of life among patients who had been diagnosed with cancer within two years.[104] The 2014 hospital care costs were inflated to 2017 using the average inflation rate for the UK.[107]

Following initial treatment discontinuation, patients are administered with subsequent lines of treatment. The possible second-line treatments include: axitinib, pazopanib, sunitinib, nivolumab, everolimus, sorafenib, bevacizumab+interferon, cabozantinib and interferon. While not all of these therapies are approved for second-line use in England, they have been included in the model since they represent the treatments given in CABOSUN and COMPARZ after disease progression. The estimated weekly drug costs of second line treatments are based on NHS list prices (Table 54 and Table 55). The distribution and duration of subsequent treatments according to initial treatment is shown in Table 56 and Table 58, respectively. In the base case, 8 weeks of waiting time was assumed based on the CABOSUN patient-level data analysis. Scenario analysis is conducted for different distribution of subsequent treatments based on opinions of UK clinicians (Table 57).

Table 54 Drug formulation, dose and total cost per week for subsequent treatments (without drug wastage)

Drug	Formulation (mg)	Cost per pack[1], £	Vials/ tabs per admin	Vials/ tabs per pack	Dose, mg	Weekly frequency	Relative dose intensity, % (SE)	Total cost per cycle, £
Axitinib	5	3517.00	1	56	10	7	102 (1.9)[2]	896.84
Pazopanib	400	1121.00	2	30	800	7	86 (8.6)[3]	449.89
Sunitinib	50	3138.84	1	28	50	4.7**	87 (6.3)[4]	457.22
Temsirolimus	30 mg/1.2 ml	620.00	0.84	1	25	1	92 (9.2)[5]	680.33
Nivolumab	40 mg/4 ml 100 mg/10 ml	439.00 1097.00	1.75 2	1 1	270*	0.5	98 (9.8)[6]	1439.00
Everolimus	10	2673.00	1	30	10	7	84 (1.1)[7]	523.28
Sorafenib	200	3576.56	2	112	800	7	80 (8.0)[8]	715.31
Cabozantinib	20/40/60	5143.00	1	30	60/20/40	7	93.3 (9.3)***	[redacted]
Bevacizumab	400 mg/16 ml 100 mg/4 ml	924.40 242.66	2 0.97	1 1	897*	0.5	88 (8.8)[8]	1037.11
Interferon	9MIU/0.5 ml	42.57	1	1	9 (MIU)	3	86 (8.6)[8]	138.00

Sources:

1 BNF, NHS indicative price. Accessed 28 November 2017[102]

2 Rini 2011. Lancet 2011.[108]

3 NICE Single technology appraisal. Pazopanib for the first-line treatment of patients with advanced renal cell carcinoma (RCC)(TA215)[18]

4 CABOSUN Clinical Study Report.[44]

5 Bevacizumab, sorafenib, tosylate, sunitinib and temsirolimus for renal cell carcinoma: A systematic review and economic evaluation. PenTAG Cost-effectiveness analysis[81]

6 NICE Single technology appraisal. Nivolumab for treated or metastatic renal cell carcinoma [ID853] [109]

7 METEOR Clinical Study Report[100]

8 Escudier et al. 2007[59]

Notes:

* Dosage according to body mass: Nivolumab-3 mg / ml, Bevacizumab-10 mg / 7kg. Average weight of 89.7 kg is taken into account (CABOSUN trial, patient-level analysis)

** Sunitinib is given in 6 weeks cycles of 4 weeks of treatment followed by a rest period of 2 weeks

*** Dose interruptions due to adverse events: 93.3% (mean).

Table 55 Drug formulation, dose and total cost per 4-weeks model cycle for subsequent treatments (with drug wastage)

Drug	Formulation (mg)	Cost per pack[1], £	Vials/ tabs per admin	Vials/ tabs per pack	Dose, mg	Weekly frequency	Relative dose intensity, % (SE)	Total cost per cycle, £
Axitinib	5	3517.00	1	56	10	7	102 (1.9)[2]	896.84
Pazopanib	400	1121.00	2	30	800	7	86 (8.6)[3]	449.89
Sunitinib	50	3138.84	1	28	50	4.7**	87 (6.3)[4]	457.22
Temsirolimus	30 mg/1.2 ml	620.00	0.84	1	25	1	92 (9.2)[5]	825.00
Nivolumab	40 mg/4 ml 100 mg/10 ml	439.00 1097.00	1.75 2	1 1	270*	0.5	98 (9.8)[6]	1572.11
Everolimus	10	2673.00	1	30	10	7	84 (1.1)[7]	523.28
Sorafenib	200	3576.56	2	112	800	7	80 (8.0)[8]	715.31
Cabozantinib	20/40/60	5143.00	1	30	60/20/40	7	93.3 (9.3)***	[redacted]
Bevacizumab	400 mg/16 ml 100 mg/4 ml	924.40 242.66	2 0.97	1 1	897*	0.5	88 (8.8)[8]	1050.02
Interferon	9MIU/0.5 ml	42.57	1	1	9 (MIU)	3	86 (8.6)[8]	155.46

Sources:

1 BNF, NHS indicative price. Assessed 28 November 2017[102]

2 Rini 2011.. Lancet 2011.[108]

3 NICE Single technology appraisal. Pazopanib for the first-line treatment of patients with advanced renal cell carcinoma (RCC)(TA215)[18]

4 CABOSUN Clinical Study Report[44]

5 Bevacizumab, sorafenib, tosylate, sunitinib and temsirolimus for renal cell carcinoma: A systematic review and economic evaluation. PenTAG Cost-effectiveness analysis[81]

6 NICE Single technology appraisal. Nivolumab for treated or metastatic renal cell carcinoma [ID853] Committee papers Committee[109]

7 METEOR Clinical Study Report [100]

8 Escudier et al. 2007[59]

Notes:

* Dosage according to body mass: Nivolumab-3mg/ml, Bevacizumab-10 mg7kg. Average weight of 89.7 kg is taken into account (CABOSUN trial, patient-level analysis)

** Sunitinib is given in 6 weeks cycles of 4 weeks of treatment followed by a rest period of 2 weeks

*** Dose interruptions due to adverse events: 93.3% (mean).

Table 56 Distribution of subsequent treatments following treatment discontinuation

	Subsequent treatments									
After progression on	**Axitinib**	**Pazo-panib**	**Sunitinib**	**Temsiro-limus**	**Nivo-lumab**	**Evero-limus**	**Sora-fenib**	**Bevaci-zumab**	**Cabo-zantinib**	**Inter-feron**
Cabozantinib [1]	23%	16%	13%	9%	13%	8%	1%	0%	1%	1%
Sunitinib [1]	19%	12%	13%	4%	15%	19%	3%	6%	6%	0%
Pazopanib [2]	6%	0%	29%	6%	0%	31%	11%	7%	0%	0%
Sources: 1 CABOSUN Clinical Study Report, Table 26.[44] 2 Pazopanib Clinical Study Report [85]										

Table 57 Distribution of subsequent treatments following treatment discontinuation (scenario analysis, cost only)

	Subsequent treatments										
After progression on	**Axiti-nib**	**Pazo-panib**	**Suniti-nib**	**Temsir-olimus**	**Nivo-lumab**	**Evero-limus**	**Sora-fenib**	**Bevaci-zumab**	**Cabo-zantinib**	**Inter-feron**	**BSC**
Cabozantinib [1]	50%	0%	0%	0%	30%	10%	0%	0%	0%	0%	10%
Sunitinib [1]	40%	0%	0%	0%	30%	10%	0%	0%	10%	0%	10%
Pazopanib [1]	40%	0%	0%	0%	30%	10%	0%	0%	10%	0%	10%
Sources: 1 Tivozanib NICE STA,[82] with the exception that best supportive care share is moved to cabozantinib. Abbreviations: BSC, best supportive care											

Table 58 Duration of subsequent treatments

Subsequent treatments	Duration (SE), weeks
Axitinib[1]	31.5 (3.2)
Pazopanib[2]	49.8 (5.0)
Sunitinib[3]	24.7 (2.5)
Temsirolimus[4]	17.0 (1.7)
Nivolumab[5]	42.0 (4.2)
Everolimus[6]	23.9 (2.4)
Sorafenib[1]	25.8 (2.6)
Bevacizumab[7]	24.0 (2.4)
Cabozantinib[6]	33.1 (3.3)
Interferon[7]	12.0 (1.2)
Sources: 1 NICE Single technology appraisal. Axitinib for the treatment of advanced renal cell carcinoma after failure of prior systemic treatment. TA333[83] 2 COMPARZ study, Clinical Study Report[85] 3 CABOSUN study, Clinical Study Report.[44] 4 Hudes et al. 2007[65] 5 NICE Single technology appraisal. Nivolumab for previously treated advanced renal cell carcinoma. TA417[33] 6 METEOR Clinical Study Report[100] 7 Rini 2008. CALGB90206 study[61]	

B.3.6 Summary of base-case analysis inputs and assumptions

Summary of base-case analysis inputs

Table 59 provides a summary of variables applied in the economic model.

Table 59 Summary of variables applied in the economic model

Variable	Value	SE	Distribution	Section
Treatment Costs				
Baseline Weight (kg)	89.70	1.5	Normal	B.3.5
Baseline age (years)	62.8	0.2	Normal	
Cost of cabozantinib (per day)	171.43	17.14	No distribution	
Cost of sunitinib (per 50 mg)	112.10	11.20	Gamma	
Cost of pazopanib (800 mg)	74.73	7.47	Gamma	
Cost of everolimus (per 10mg)	89.10	8.91	Gamma	
Cost of axitinib (per 10 mg)	125.61	12.56	Gamma	
Cost of nivolumab (per 3 mg)	32.91	3.29	Gamma	
Cost of interferon (per 9 MIU)	42.57	4.26	Gamma	
Cost of bevacizumab (per 100 mg)	242.66	24.27	Gamma	
Cost of bevacizumab (per 400 mg)	924.40	92.44	Gamma	
Cost of temsirolimus (per 25 mg)	516.67	51.67	Gamma	
Cost of sorafenib (per 800mg)	127.73	12.77	Gamma	
Relative dose intensity of cabozantinib for 1st line	0.943	0.015	Beta	
Relative dose intensity of cabozantinib for 2nd line	0.933	0.013	Beta	
Relative dose intensity of sunitinib	0.87	0.06	Beta	
Relative dose intensity of pazopanib	0.86	0.09	Beta	
Relative dose intensity of everolimus	0.84	0.01	Beta	
Relative dose intensity of axitinib	1.02	0.02	Beta	
Relative dose intensity of interferon alone	0.86	0.09	Beta	
Relative dose intensity of bevacizumab	0.88	0.09	Beta	
Relative dose intensity of temsirolimus	0.92	0.09	Beta	
Relative dose intensity of nivolumab	0.98	0.10	Beta	
Relative dose intensity of sorafenib	0.80	0.08	Beta	
Wastage of nivolumab per administration	0.08	-	Beta	
Single administration cost of interferon	9.5	0.95	Gamma	

Single administration cost of bevacizumab	199	19.9	Gamma	
Single administration cost of temsirolimus	199	19.9	Gamma	
Single administration cost of nivolumab	199	19.9	Gamma	
AE costs				
Total cost of cabozantinib AE	354.02	35.40	Gamma	B.3.5
Total cost of sunitinib AE	256.89	25.89	Gamma	
Total cost of pazopanib AE	162.53	16.25	Gamma	
Disease Management costs				
Cost of the initial consultation	219	21.9	Gamma	B.3.5
Cost of follow up consultation	173	17.3	Gamma	
Cost of CT scan	115	11.5	Gamma	
Cost of Blood test	3	0.3	Gamma	
Follow up consultation frequency per week - pre progression	0.25	-	No distribution	
CT scan frequency per week -pre progression	0.08	-	No distribution	
Blood test frequency per week - pre progression	0.25	-	No distribution	
Follow up consultation frequency per week - progressed disease	0.25	-	No distribution	
CT scan frequency per week - progressed disease	0.08	-	No distribution	
Blood test frequency per week - progressed disease	0.25	-	No distribution	
End-of-life costs				
End-of-life cost	6207.59	620.76	Gamma	B.3.5
Time on follow up treatment				
Time to second line treatment (weeks)	8	0.8	Normal	B.3.5
Time on axitinib 2nd line (weeks)	31.5	3.15	Normal	
Time on pazopanib 2nd line (weeks)	49.8	4.98	Normal	
Time on sunitinib 2nd line (weeks)	24.7	2.47	Normal	
Time on temsirolimus 2nd line (weeks)	17.0	1.70	Normal	
Time on nivolumab 2nd line (weeks)	42.0	4.20	Normal	
Time on everolimus 2nd line (weeks)	23.9	2.39	Normal	
Time on sorafenib 2nd line (weeks)	25.8	2.58	Normal	
Time on bevacizumab 2nd line (weeks)	24.0	2.40	Normal	
Time on cabozantinib 2nd line (weeks)	33.1	3.31	Normal	
Time on interferon 2nd line (weeks)	12.0	1.20	Normal	

Utilities				
Utilities: Progression-free state	0.726	0.011	Beta	B.3.5
Utilities: Progressed disease	0.649	0.019	Beta	
AE duration (weeks)	4.00	-	No distribution	
AE average episodes per patient	1.43	-	No distribution	
Utility decrement due to Grade 3/4 AEs	-0.20	-0.07	Beta	
Total Disutilities of AE for cabozantinib	-0.023		No distribution	
Total Disutilities of AE for sunitinib	-0.022		No distribution	
Total Disutilities of AE for pazopanib	-0.020		No distribution	
CABOSUN data, PFS, lognormal distribution, cabo	Intercept 2.295 Scale 1.073	Intercept 0.145 Scale 0.121	Multivariate normal distribution	
CABOSUN data, PFS, lognormal distribution, suni	Intercept 1.637 Scale 0.915	Intercept 0.122 Scale 0.093	Multivariate normal distribution	
CABOSUN data, OS, exponential distribution, cabo	Intercept 3.657	Intercept 0.162	Normal distribution	
CABOSUN data, OS, exponential distribution, suni	Intercept 3.380	Intercept 0.149	Normal distribution	
CABOSUN data, TTD, lognormal distribution, cabo	Intercept 1.872 Scale 1.184	Intercept 0.136 Scale 0.105	Multivariate normal distribution	
ITC, fractional polynomial curve, PFS, P1=-1 and P2=-1, cabo	Para1 -4.099 Para2 -1.414 Para3 6.173	Para1 0.598 Para2 1.272 Para3 2.377	Multivariate normal distribution	
ITC, fractional polynomial curve, PFS, P1=-1 and P2=-1, suni	Para1 -2.822 Para2 -0.148 Para3 3.118	Para1 0.552 Para2 0.597 Para3 1.490	Multivariate normal distribution	
ITC, fractional polynomial curve, PFS, P1=-1 and P2=-1, pazo	Para1 -2.714	Para1 0.563 Para2	Multivariate normal distribution	

	Para2 -0.858 Para3 3.689	0.642 Para3 1.532		
ITC, fractional polynomial curve, OS, P1=-1 and P2=-1, cabo	Para1 -3.928 Para2 -3.463 Para3 3.006	Para1 0.518 Para2 2.350 Para3 2.833	Multivariate normal distribution	
ITC, fractional polynomial curve, OS, P1=-1 and P2=-1, suni	Para1 -3.492 Para2 -0.923 Para3 1.068	Para1 0.364 Para2 0.741 Para3 1.294	Multivariate normal distribution	
ITC, fractional polynomial curve, OS, P1=-1 and P2=-1, pazo	Para1 -3.359 Para2 -0.769 Para3 0.073	Para1 0.422 Para2 1.103 Para3 1.903	Multivariate normal distribution	
Abbreviations: AE, adverse event; CI, confidence interval; GP, general practitioner; HR, hazard ratio; ITC, indirect treatment comparison; ITT, intent-to-treat; OS, overall survival; PFS, progression-free survival; PPS, post-progression survival				

Assumptions

The base case analysis is subject to several key assumptions, described and discussed throughout Section 5. The key assumptions are:

Effectiveness

- OS and PFS curves for cabozantinib and sunitinib fitted to the CABOSUN patient-level data are best represented by exponential and log-normal curves, respectively. Comparators modelled by using the ITC are best represented by FP 2nd order model (P1=-1 and P2=-1) fitted to re-generated patient-level data of comparator studies (COMPARZ).
- The relative efficacy for cabozantinib vs. pazopanib in the model is based on the fractional polynomial ITC, which assumes that there are no significant imbalances in effect modifiers between different types of direct comparisons.

- Given that there are differences in baseline risk group, this assumption is discussed. It was not possible to re-run the ITC for particular subgroups due to lack of data.

Quality of life

- Quality of life is dependent on disease progression status and toxicity of treatments.
- The most suitable source to estimate utilities are the data from the recent tivozanib NICE appraisal for all comparators, to avoid combining several sources/methods of preference elicitation together.

Resource use and costs

- Treatment duration is best characterised by log-normal curve for cabozantinib and sunitinib. No TTD was identified for pazopanib and the median treatment duration and median PFS were used to estimate pazopanib TTD.
- Vial sharing will not occur in practice in the administration of nivolumab, bevacizumab, temsirolimus and interferon. Hence drug wastage is assumed.
- Resource use and costs for disease management are dependent on RECIST-defined progression status.
- Management of grade 3 and 4 AEs are associated with resource use validated by UK oncologists.

B.3.7 Base-case results

Base-case incremental cost-effectiveness analysis results

Base case results of pairwise comparison are shown in Table 60 and Table 61. Cabozantinib has proven to be an effective treatment for advanced RCC in treatment naïve patients when compared with sunitinib, with a predicted survival benefit of 0.657 years (0.401 QALYs). Cabozantinib is also an effective treatment compared with pazopanib.

Table 60 Base-case results: pair wise analysis of cabozantinib versus sunitinib (from CABOSUN study)

Drug	Total costs	Total QALYs	Total life-years	ΔCosts	ΔQALYs	ΔLife years	ICER vs baseline	ICER
Sunitinib	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	15,170	0.401	0.657	37,793	37,793
Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years								

Table 61 Base-case results: pair wise analysis of cabozantinib versus comparators (based on the ITC results)

Drug	Total costs	Total QALYs	Total life-years	ΔCosts	ΔQALYs	ΔLife years	ICER vs baseline	ICER
Pazopanib	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Sunitinib	[redacted]	[redacted]	[redacted]	7,561	-0.021	-0.035	Dominated by pazopanib	Dominated by pazopanib
Cabozantinib	[redacted]	[redacted]	[redacted]	23,526	0.486	0.799	48,451	31,538

Clinical outcomes from the model and disaggregated results of the base case incremental cost-effectiveness analysis are provided in Appendix J.

B.3.8 Sensitivity analyses

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed to translate the imprecision in all input variables into a measure of decision uncertainty in the cost-effectiveness model for the options being compared. The point estimates and distribution choices are described for each parameter in Table 59. Uncertainties for distributions derived from the ITC were tested by sampling using variance-covariance matrix and random draws from the multivariate-normal distribution. Table 62, Table 63 and Table 64 report the mean probabilistic base case results for all pairwise comparisons for cabozantinib. Figure 17, Figure 18 and Figure 19 show a PSA scatterplot and Figure 20, Figure 21 and Figure 22 show the cost-acceptability curve for cabozantinib versus key comparators for 1,000 PSA iterations. Comparisons versus sunitinib are CABOSUN- and ITC-based and comparisons versus pazopanib are ITC-based. Scatterplots show that there is some parameter uncertainty around the mean ICER.

The results of CABOSUN-based and ITC-based analyses suggest that the probability of cabozantinib being cost-effective versus sunitinib at a willingness-to pay threshold of £50,000 per QALY gained is 66.1% and 74.4%, respectively . The results suggest that the probability of cabozantinib being cost-effective versus pazopanib at a willingness-to pay threshold of £50,000 per QALY gained is 47.8%.

Every effort has been made to ensure that parameter uncertainty was informed by data and not arbitrary assumptions for key parameters.

Table 62 Mean probabilistic base case results for cabozantinib vs. sunitinib (CABOSUN-based)

				Incremental to drug			
	Total Costs	**QALYs**	**Life-Years**	**Costs**	**QALYs**	**Life Years**	**ICER**
Sunitinib	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	14,722	0.392	0.645	37,592
Abbreviations: ICER, incremental cost-effectiveness ratio, QALY, quality-adjusted life year							

Table 63 Mean probabilistic base case results for cabozantinib vs. sunitinib (ITC-based)

				Incremental to drug			
	Total Costs	**QALYs**	**Life-Years**	**Costs**	**QALYs**	**Life Years**	**ICER**
Sunitinib	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	16,673	0.551	0.943	30,239
Abbreviations: ICER, incremental cost-effectiveness ratio, QALY, quality-adjusted life year							

Table 64 Mean probabilistic base case results for cabozantinib vs. pazopanib (ITC-based)

				Incremental to drug			
	Total Costs	**QALYs**	**Life-Years**	**Costs**	**QALYs**	**Life Years**	**ICER**
Pazopanib	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	24,191	0.504	0.839	48,030
Abbreviations: ICER, incremental cost-effectiveness ratio, QALY, quality-adjusted life year							

Figure 17 PSA scatter plot cabozantinib vs. sunitinib (CABOSUN- based)

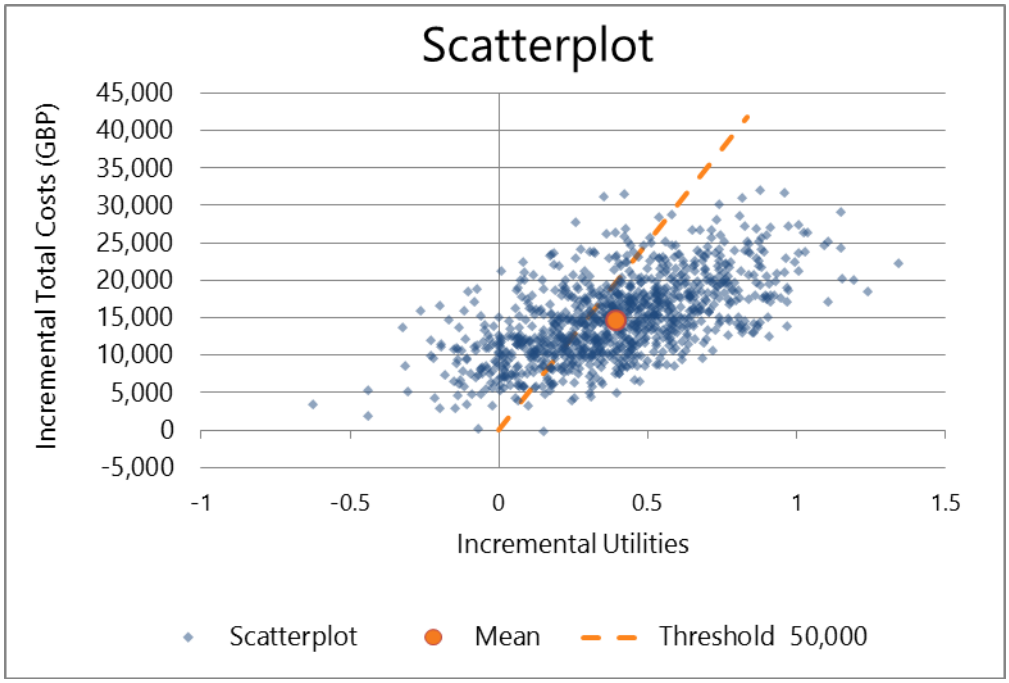


Figure 18 PSA scatter plot cabozantinib vs. sunitinib (ITC-based)

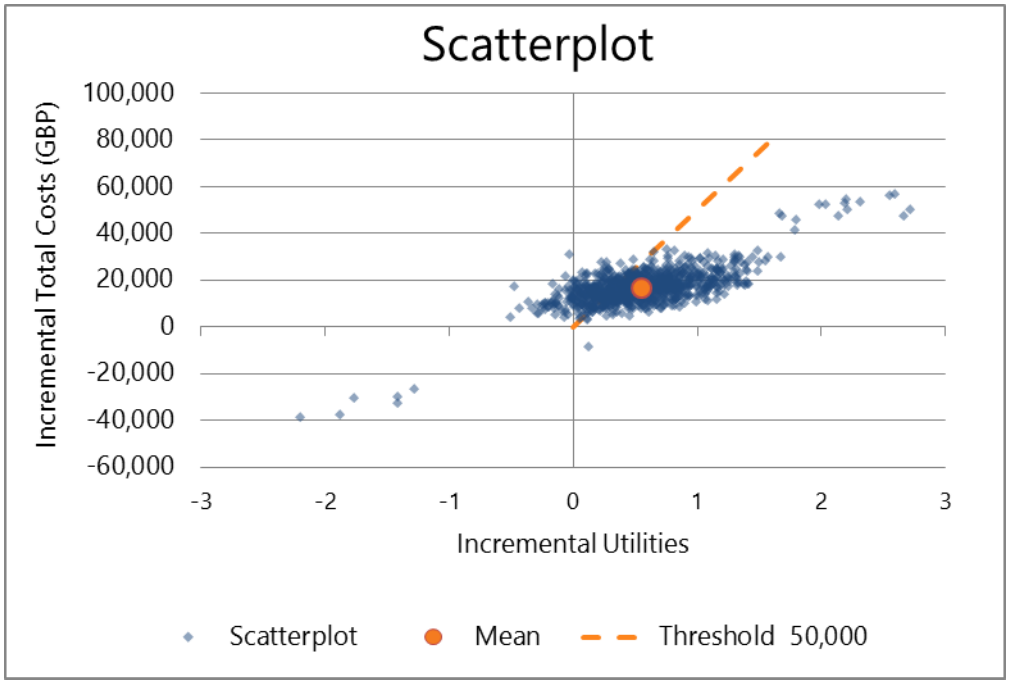


Figure 19 PSA scatter plot cabozantinib vs. pazopanib (ITC-based)

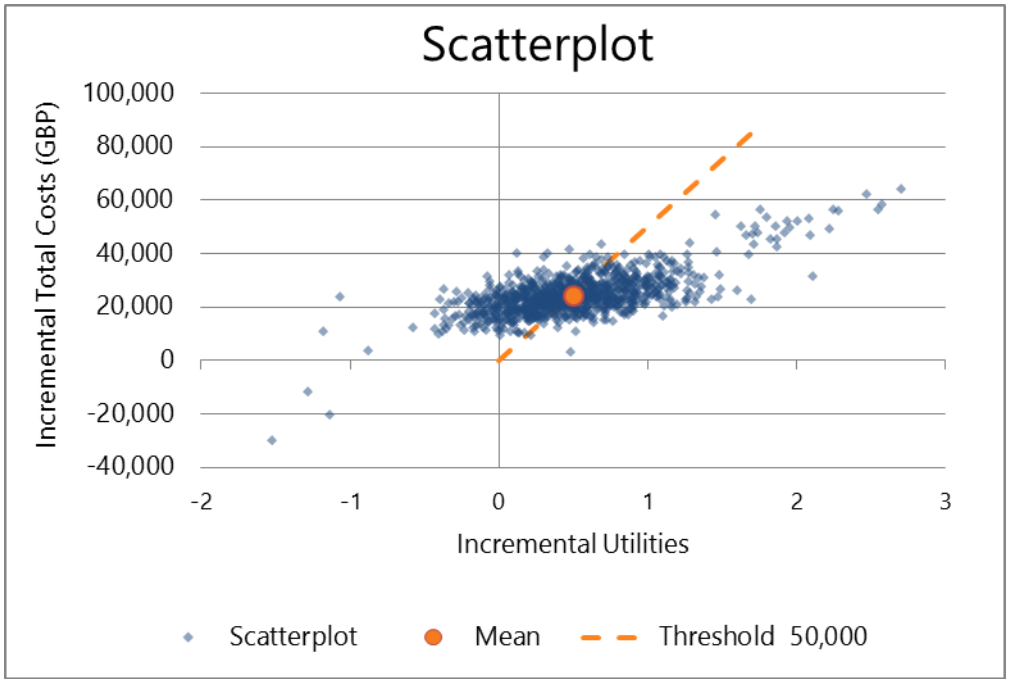


Figure 20 PSA cost-effectivness acceptability curve cabozantinib vs sunitinib (CABOSUN-based)

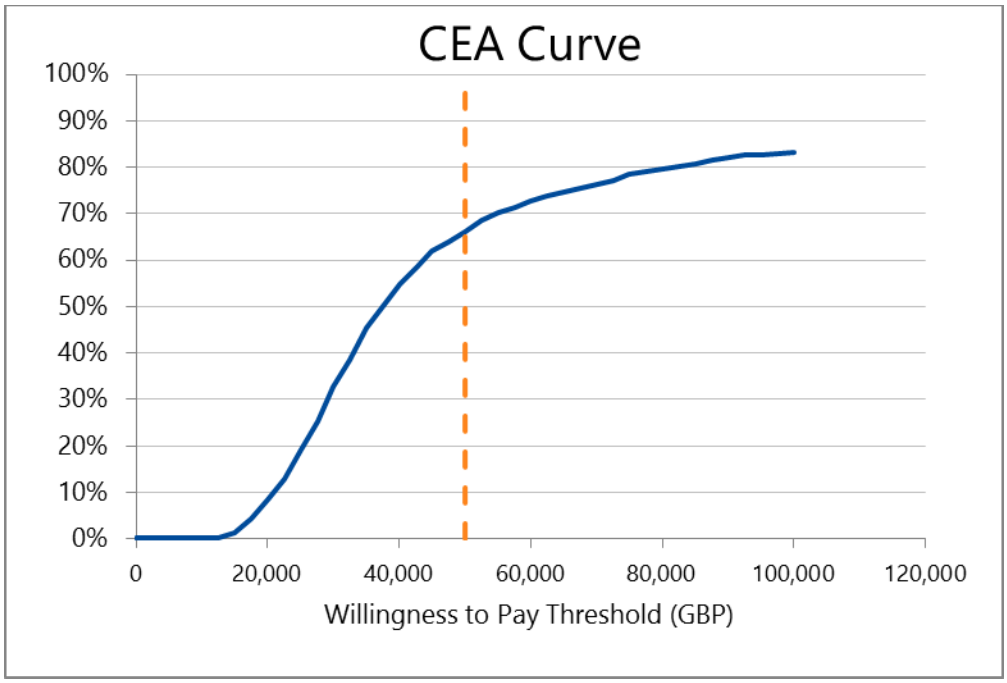


Figure 21 PSA cost-effectivness acceptability curve cabozantinib vs sunitinib (ITC-based)

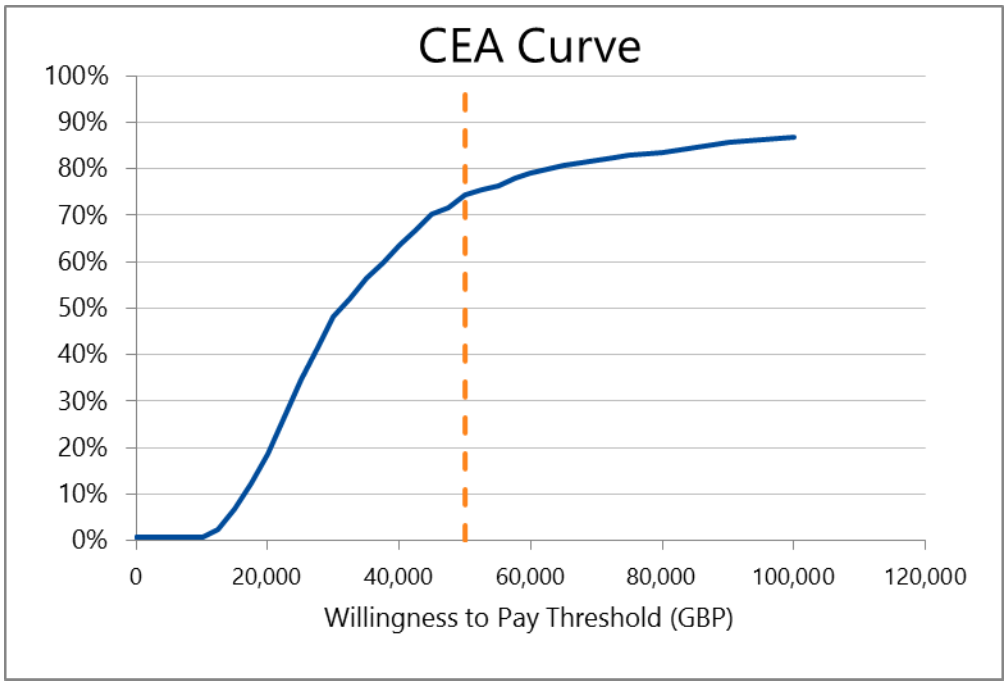
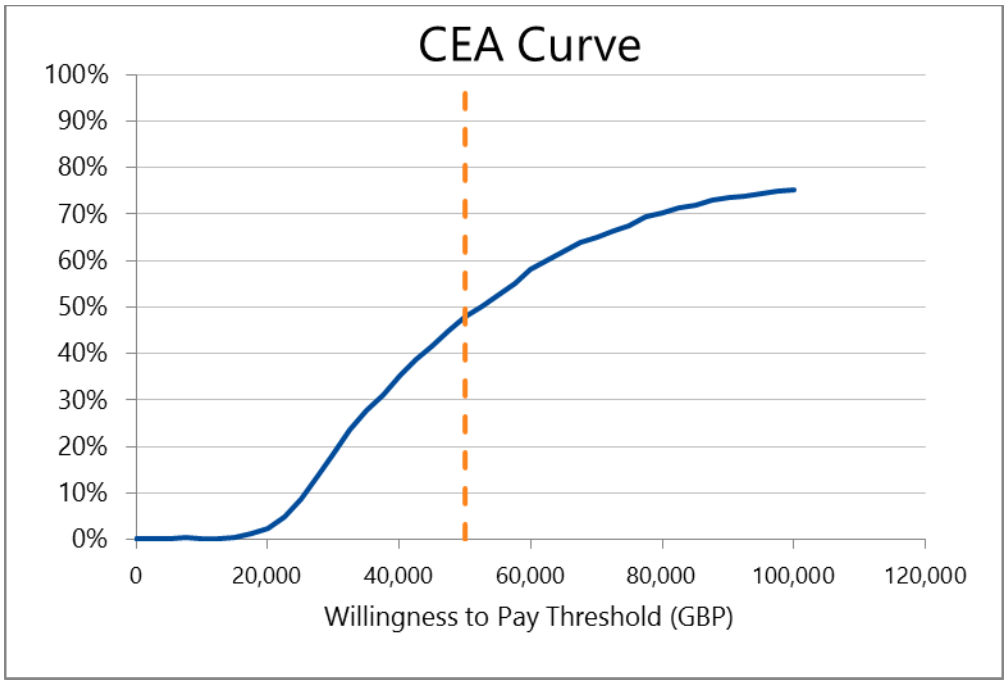


Figure 22 PSA cost-effectiveness acceptability curve cabozantinib vs pazopanib (ITC-based)



Deterministic sensitivity analysis

An assessment of parameter uncertainty was also performed via deterministic sensitivity analysis (DSA). The model parameter values were individually varied to test the sensitivity of the model's results to specific parameters or sets of parameters. The inputs and the range tested are reported in Table 65.

Table 65 Summary of variables included in DSA

Variable	Base case	Min, max
General		
Time horizon (10 - 20 years)	20	10;20
Cost Discount (%)	0.035	0;0.06
Effect Discount (%)	0.035	0;0.06
Utilities		
Utilities: Progression-free state	0.726	Beta
Utilities: Progressed state	0.649	Beta
Adverse event-related		
Disutility values for AE	0.204	Normal
Total AE cost for Cabozantinib	339	Gamma
Total AE cost for Sunitinib	254	Gamma
Total AE cost for Pazopanib	159	Gamma
Demographics		
Baseline Weight (kg)	89.7	Normal
Baseline age (years)	62.8	Normal
Proportion of male at baseline (%)	78%	Beta
Drug and administration costs		
Relative dose intensity of cabozantinib for 1st line	94.3%	Beta
Relative dose intensity of cabozantinib for 2nd line	93.3%	Beta
Cost of sunitinib (Sutent®) (50 mg), daily	112.1	Gamma
Relative dose intensity of sunitinib	87%	Beta
Cost of pazopanib (Votrient®) (800 mg), daily	74.7	Gamma
Relative dose intensity of pazopanib	86%	Beta
Cost of everolimus (Afinitor®) (10 mg)	89.1	Gamma
Relative dose intensity of everolimus	84%	Beta
Cost of axitinib (Inlyta®) (10 mg)	125.6	Gamma
Relative dose intensity of axitinib	102%	Normal
Cost of nivolumab (Opdivo) (3 mg)	32.9	Gamma
Relative dose intensity of nivolumab	98%	Beta
Wastage of nivolumab per administration	8%	Beta
Single administration cost of nivolumab	199.0	Gamma
Cost of interferon (Pegasys®) (9 million units)	42.6	Gamma

Relative dose intensity of interferon	86%	Beta
Injection cost (district nurse visit)	9.5	Gamma
Cost of bevacizumab (Avastin®) (100 mg)	242.7	Gamma
Cost of bevacizumab (Avastin®) (400 mg)	924.4	Gamma
Relative dose intensity of bevacizumab	88%	Beta
Administration cost for bevacizumab	199.0	Gamma
Cost of temsirolimus (Torisel®) (25 mg)	516.7	Gamma
Relative dose intensity of temsirolimus	92%	Beta
Administration cost for temsirolimus	199.0	Gamma
Cost of sorafenib (Nexavar) (800 mg)	127.7	Gamma
Relative dose intensity of sorafenib	80%	Beta
Disease management costs		
Cost of outpatient consultation (initial)	197	Gamma
Cost of outpatient consultation (follow-up)	163	Gamma
Cost of CT scan	140	Gamma
Cost of blood test	3	Gamma
PFS: outpatient (follow-up) frequency	4	No distribution
PFS: CT scan frequency	12	No distribution
PFS: blood test frequency	4	No distribution
Progression: outpatient (follow-up) frequency	4	No distribution
Progression: CT scan frequency	12	No distribution
Progression: blood test frequency	4	No distribution
End-of-life costs		
End-of-life cost	6,208	Gamma
Subsequent treatment proportion		
After progression of Cabozantinib 2nd line: Axitinib	23%	No distribution
After progression of Cabozantinib 2nd line: Pazopanib	16%	No distribution
After progression of Cabozantinib 2nd line: Sunitinib	13%	No distribution
After progression of Cabozantinib 2nd line: Temsirolimus	9%	No distribution
After progression of Cabozantinib 2nd line: Nivolumab	13%	No distribution
After progression of Cabozantinib 2nd line: Everolimus	8%	No distribution
After progression of Cabozantinib 2nd line: Sorafenib	1%	No distribution
After progression of Cabozantinib 2nd line: Bevacizumab	0%	No distribution
After progression of Cabozantinib 2nd line: Cabozantinib	1%	No distribution
After progression of Cabozantinib 2nd line: Interferon	1%	No distribution
After progression of Sunitinib 2nd line: Axitinib	19%	No distribution
After progression of Sunitinib 2nd line: Pazopanib	12%	No distribution
After progression of Sunitinib 2nd line: Sunitinib	13%	No distribution
After progression of Sunitinib 2nd line: Temsirolimus	4%	No distribution
After progression of Sunitinib 2nd line: Nivolumab	15%	No distribution

After progression of Sunitinib 2nd line: Everolimus	19%	No distribution
After progression of Sunitinib 2nd line: Sorafenib	3%	No distribution
After progression of Sunitinib 2nd line: Bevacizumab	6%	No distribution
After progression of Sunitinib 2nd line: Cabozantinib	6%	No distribution
After progression of Sunitinib 2nd line: Interferon	0%	No distribution
After progression of Pazopanib 2nd line: Axitinib	6%	No distribution
After progression of Pazopanib 2nd line: Pazopanib	0%	No distribution
After progression of Pazopanib 2nd line: Sunitinib	29%	No distribution
After progression of Pazopanib 2nd line: Temsirolimus	6%	No distribution
After progression of Pazopanib 2nd line: Nivolumab	0%	No distribution
After progression of Pazopanib 2nd line: Everolimus	31%	No distribution
After progression of Pazopanib 2nd line: Sorafenib	11%	No distribution
After progression of Pazopanib 2nd line: Bevacizumab	7%	No distribution
After progression of Pazopanib 2nd line: Cabozantinib	0%	No distribution
After progression of Pazopanib 2nd line: Interferon	0%	No distribution
Treatment time on the 2nd line		
Time to 2nd line treatment (weeks)	8	Normal
Time on 2nd line treatment: Axitinib (weeks)	31.5	Normal
Time on 2nd line treatment: Pazopanib (weeks)	49.8	Normal
Time on 2nd line treatment: Sunitinib (weeks)	24.7	Normal
Time on 2nd line treatment: Temsirolimus (weeks)	17.0	Normal
Time on 2nd line treatment: Nivolumab (weeks)	42.0	Normal
Time on 2nd line treatment: Everolimus (weeks)	23.9	Normal
Time on 2nd line treatment: Sorafenib (weeks)	25.8	Normal
Time on 2nd line treatment: Bevacizumab (weeks)	24.0	Normal
Time on 2nd line treatment: Cabozantinib (weeks)	33.1	Normal
Time on 2nd line treatment: Interferon (weeks)	12.0	Normal
CABOSUN data, PFS, lognormal distribution, cabo	Intercept 2.295 Scale 1.073	Multivariate normal distribution
CABOSUN data, PFS, lognormal distribution, suni	Intercept 1.637 Scale 0.915	Multivariate normal distribution
CABOSUN data, OS, exponential distribution, cabo	Intercept 3.657	Normal distribution
CABOSUN data, OS, exponential distribution, suni	Intercept 3.380	Normal distribution
CABOSUN data, TTD, lognormal distribution, cabo	Intercept 1.872 Scale 1.184	Multivariate normal distribution
ITC, fractional polynomial curve, PFS, P1=-1 and P2=-1, cabo	Para1 -4.099 Para2	Multivariate normal distribution

	-1.414 Para3 6.173	
ITC, fractional polynomial curve, PFS, P1=-1 and P2=-1, suni	Para1 -2.822 Para2 -0.148 Para3 3.118	Multivariate normal distribution
ITC, fractional polynomial curve, PFS, P1=-1 and P2=-1, pazo	Para1 -2.714 Para2 -0.858 Para3 3.689	Multivariate normal distribution
ITC, fractional polynomial curve, OS, P1=-1 and P2=-1, cabo	Para1 -3.928 Para2 -3.463 Para3 3.006	Multivariate normal distribution
ITC, fractional polynomial curve, OS, P1=-1 and P2=-1, suni	Para1 -3.492 Para2 -0.923 Para3 1.068	Multivariate normal distribution
ITC, fractional polynomial curve, OS, P1=-1 and P2=-1, pazo	Para1 -3.359 Para2 -0.769 Para3 0.073	Multivariate normal distribution

Abbreviations: AE, adverse event; CI, confidence interval; GP, general practitioner; HR, hazard ratio; ITC, indirect treatment comparison; ITT, intention-to-treat; OS, overall survival; Pazo, pazopanib; PFS, progression-free survival; PPS, post-progression survival; Suni, sunitinib

Figure 23 and Figure 24 show tornado diagrams depicting the parameters which change the ICER by more than £1,000/QALY compared to the base case.

Figure 23 Tornado graph, cabozantinib vs sunitinib

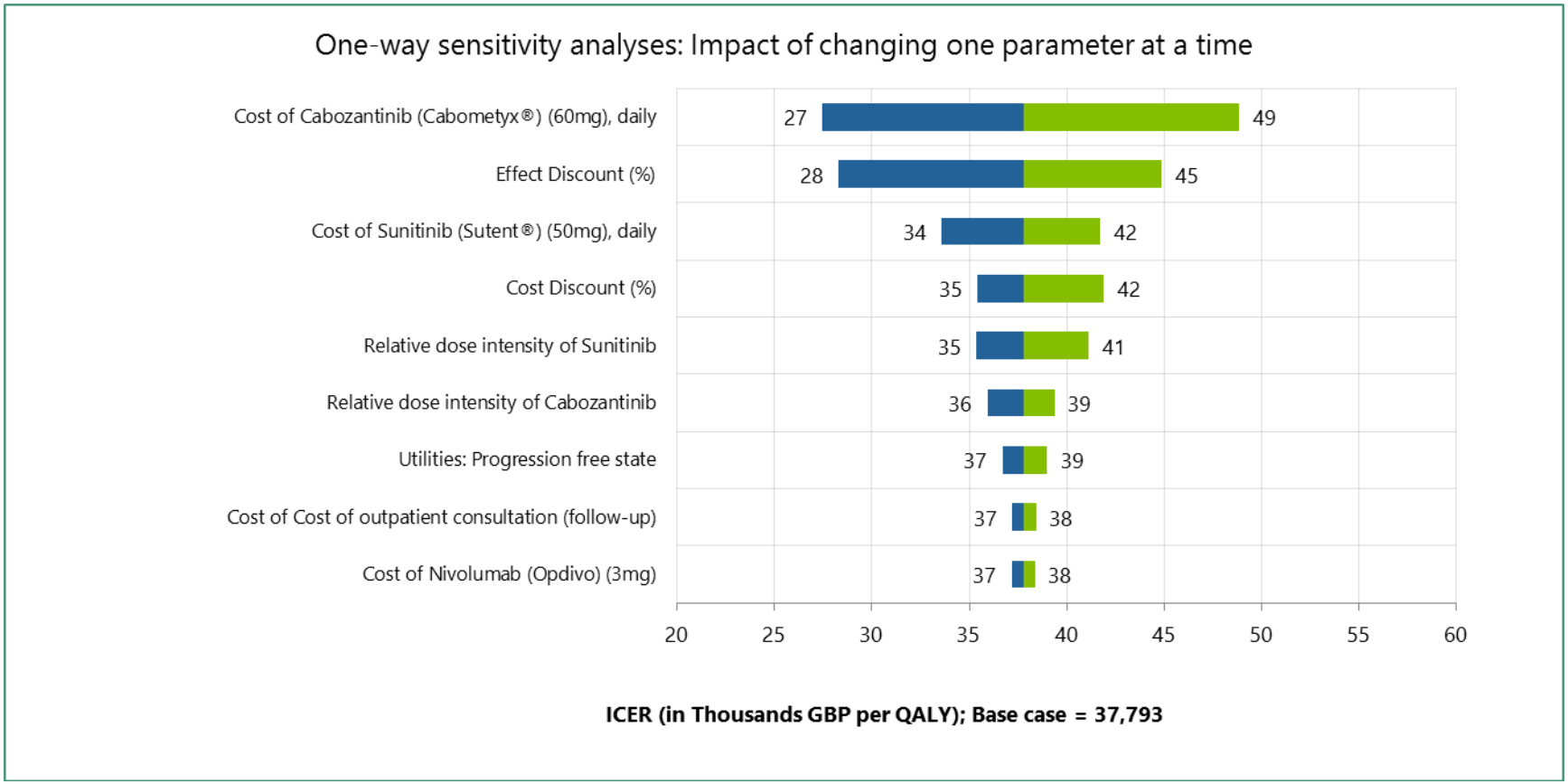
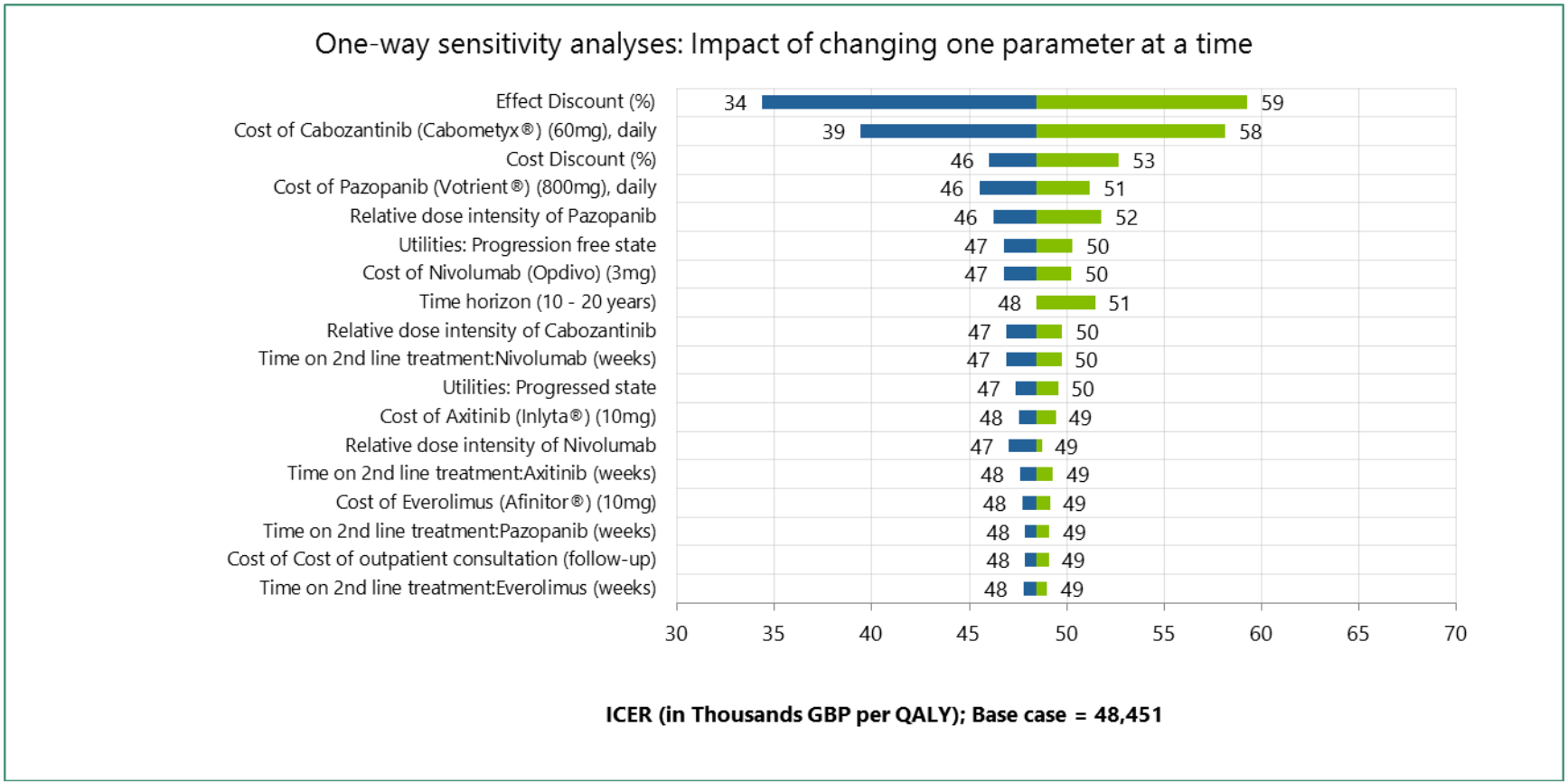


Figure 24 Tornado graph, cabozantinib vs pazopanib



Scenario analysis

Uncertainty in a parameter may be represented by several discrete values, instead of a continuous range, and these are tested in the scenario analyses. The scenarios tested and the results are shown in Table 66 and Table 67.

Table 66 Scenario analysis (cabozantinib vs sunitinib)

Scenario		Total cost		Total QALY		ICER
		Cabo.	Suni	Cabo.	Suni	
Base case		[redacted]	[redacted]	[redacted]	[redacted]	37,793
Discount	0%	[redacted]	[redacted]	[redacted]	[redacted]	31,368
	6%	[redacted]	[redacted]	[redacted]	[redacted]	42,053
Time horizon	10 years	[redacted]	[redacted]	[redacted]	[redacted]	38,360
CABOSUN data						
PFS curves	PFS= exponential	[redacted]	[redacted]	[redacted]	[redacted]	35,910
	PFS= Weibull	[redacted]	[redacted]	[redacted]	[redacted]	34,537
	PFS= Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	37,270
OS curves	OS= exponential	[redacted]	[redacted]	[redacted]	[redacted]	37,793
	OS= Weibull	[redacted]	[redacted]	[redacted]	[redacted]	46,660
	OS= Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	34,895
ITC: parametric curve method, random model						
PFS and OS curves	PFS = OS = exponential	[redacted]	[redacted]	[redacted]	[redacted]	34,401
	PFS = OS = Weibull	[redacted]	[redacted]	[redacted]	[redacted]	41,250
	PFS = OS = Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	28,736
ITC: fractional polynomial method						
OS curves	2nd order model (P1=-0.5, P2=0)	[redacted]	[redacted]	[redacted]	[redacted]	55,541
	2nd order model (P1= -1, P2=0)	[redacted]	[redacted]	[redacted]	[redacted]	38,735
TTD curves	TTD= exponential	[redacted]	[redacted]	[redacted]	[redacted]	31,458
	TTD= Weibull	[redacted]	[redacted]	[redacted]	[redacted]	31,470
	TTD= Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	34,285
	TTD= gamma	[redacted]	[redacted]	[redacted]	[redacted]	35,345
Utility values	Swinburn	[redacted]	[redacted]	[redacted]	[redacted]	31,196

	Pazo NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	38,727
	Suni NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	35,985
Age-adjusted utilities	Exclude	[redacted]	[redacted]	[redacted]	[redacted]	36,171
AE disutility source	METEOR AE disutilities	[redacted]	[redacted]	[redacted]	[redacted]	37,792
Cost	Wastage excluded	[redacted]	[redacted]	[redacted]	[redacted]	37,961
	Subsequent treatment cost (UK clinicians)	[redacted]	[redacted]	[redacted]	[redacted]	40,281
	Blood test (comprehensive test)	[redacted]	[redacted]	[redacted]	[redacted]	38,156
	Health resource (UK clinicians)	[redacted]	[redacted]	[redacted]	[redacted]	37,793
	End-of-life cost excluded	[redacted]	[redacted]	[redacted]	[redacted]	38,153
Abbreviations: Cabo, cabozantinib; Pazo, pazopanib; PFS, progression-free survival; OS, overall survival, Suni, sunitinib; TTD, time to discontinuation						

Table 67 Scenario analysis (cabozantinib vs pazopanib)

Scenario		Total cost		Total QALY		ICER
		Cabo.	Pazo	Cabo.	Pazo	
Base case		[redacted]	[redacted]	[redacted]	[redacted]	48,451
Discount	0%	[redacted]	[redacted]	[redacted]	[redacted]	37,382
	6%	[redacted]	[redacted]	[redacted]	[redacted]	56,288
Time horizon	10 years	[redacted]	[redacted]	[redacted]	[redacted]	51,450
ITC: parametric curve method, random model						
PFS and OS curves	PFS = OS = exponential	[redacted]	[redacted]	[redacted]	[redacted]	58,512
	PFS = OS = Weibull	[redacted]	[redacted]	[redacted]	[redacted]	71,510
	PFS = OS = Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	45,255
ITC: fractional polynomial method						
OS curves	2nd order model (P1=-0.5, P2=0)	[redacted]	[redacted]	[redacted]	[redacted]	81,252
	2nd order model (P1= -1, P2=0)	[redacted]	[redacted]	[redacted]	[redacted]	58,566

TTD curves	TTD= exponential	[redacted]	[redacted]	[redacted]	[redacted]	42,967
	TTD= Weibull	[redacted]	[redacted]	[redacted]	[redacted]	42,908
	TTD= Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	45,440
	TTD= gamma	[redacted]	[redacted]	[redacted]	[redacted]	46,107
Utility values	Swinburn	[redacted]	[redacted]	[redacted]	[redacted]	35,870
	Pazo NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	48,946
	Suni NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	46,684
Age-adjusted utilities	Exclude	[redacted]	[redacted]	[redacted]	[redacted]	46,028
AE disutility source	METEOR AE disutilities	[redacted]	[redacted]	[redacted]	[redacted]	48,446
Cost	Wastage excluded	[redacted]	[redacted]	[redacted]	[redacted]	47,515
	Subsequent treatment cost (UK clinicians)	[redacted]	[redacted]	[redacted]	[redacted]	32,474
	Blood test (comprehensive test)	[redacted]	[redacted]	[redacted]	[redacted]	48,815
	Health resource (UK clinicians)	[redacted]	[redacted]	[redacted]	[redacted]	48,451
	End-of-life cost excluded	[redacted]	[redacted]	[redacted]	[redacted]	48,828
Abbreviations: Cabo, cabozantinib; ICER, incremental cost-effectiveness ratio, Pazo, pazopanib; PFS, progression-free survival; OS, overall survival, Suni, sunitinib; QALY, quality-adjusted life year						

Summary of sensitivity analyses results

The results of the one-way sensitivity analyses demonstrated that the model was most sensitive to changes in: drug cost (cost of cabozantinib) and the effect of discounting. Other inputs had little impact on the model results. The scenario analyses also showed the robustness of the base case results. The analysis that had the biggest impact on the results was the use of the OS curve FP 2nd order model (P1=-0.5, P2=0), although clinicians did not think that this was a clinically plausible fit. Using this curve resulted in sunitinib patients having lower survival than patients receiving pazopanib (see Appendix J).

The results of the PSA demonstrated that cabozantinib had a 66.1% and 74.4% probability of being cost-effective versus sunitinib at a threshold of £50,000 based on CABOSUN and ITC analyses, respectively. Furthermore, cabozantinib had a 47.8% probability of being cost-effective versus pazopanib at a threshold of £50,000. The key model drivers were identified from different sensitivity analyses: cost-related model parameters and the choice of curve type for OS parametric modelling.

B.3.9 Subgroup analysis

Subgroup analyses of clinical outcomes (shown in Appendix E) showed the estimated clinical benefits of cabozantinib versus sunitinib from CABOSUN data to be consistent across a range of subgroups. Subgroup analyses were not explored in the economic model.

B.3.10 Validation

Validation of cost-effectiveness analysis

Previous appraisals of advanced and metastatic RCC treatments were reviewed and the economic approach taken for cabozantinib within this submission was designed to be consistent. The aim was to analyse key clinical outcomes that impact NHS/PSS costs and patients' HRQL.

The clinical outputs of the model were validated with UK clinical oncologists, and the model validation was carried out by economists who were not involved in the

development of the original economic model. The list below gives an overview of the validation routines carried out:

- Input data validation
  - Rationale for inclusion of particular data sources
  - Data sources checked against original source
  - Distributions and parameters to represent uncertainty
  - Data adjustments:
    - ◊ Mathematical transformations, treatment of outliers, treatment of missing data, data synthesis, calibration, etc.
- Technical validation
  - Detection of coding errors
  - Sheet by sheet testing, including macros
  - Check formulas on each input cell and how the linking of data to the variables/engines is done.
  - Check model parameters, testing of dropdown menus, names of cells, and all switches, including all sensitivity analyses
  - Check if any elements seem redundant
  - Check intended functionality of macros versus actual functionality, and for interpretability
  - Run model with extreme values
  - Movement of patients through the model
  - Additional checks:
    - ◊ Suggestions for optimisation for speed and accuracy, if relevant
    - ◊ Absence of bugs
    - ◊ Logical code structure
    - ◊ Appropriate transition of the conceptual model
    - ◊ Appropriateness of data and model

B.3.11 Interpretation and conclusions of economic evidence

Cabozantinib delivers clinically meaningful improvements in OS and PFS, while maintaining a manageable toxicity profile. In treatment naive intermediate and poor risk groups of advanced RCC patients, treatment with cabozantinib was more costly but also more effective in terms of LYs and QALYs gained than treatment with sunitinib or pazopanib. Specifically, cabozantinib yielded an overall ICER of £23,095/LY gained and £37,793/QALY compared to sunitinib, the current standard of care. Cabozantinib also extended life by 0.66 LYs and provided 0.401 incremental QALYs gained compared to sunitinib. Compared to pazopanib, the overall ICER was £29,462/LY gained and £48,451/QALY. Cabozantinib also extended life by 0.80 LYs and provided 0.49 incremental QALYs gained compared to pazopanib. The analysis was driven by the difference in PFS and OS between cabozantinib and sunitinib, as well as the difference in treatment costs.

The results of the one-way sensitivity analyses demonstrated that the model was most sensitive to changes in: drug cost (cost of cabozantinib) and the effect of discounting. Other inputs had less impact on the model results. The scenario analyses also showed the robustness of the base case results. The results of the PSA demonstrated that cabozantinib had a 66.1%, 74.4% and 47.8% probability of being cost-effective at a threshold of £50,000 versus sunitinib, based on CABOSUN and ITC results, and pazopanib, respectively. The key model drivers were identified from different sensitivity analyses: cost related model parameters and the choice of curve type for OS parametric modelling.

The key strength of this analysis was that it was based on evidence from an ITC comparing parametric survival curves, rather than HRs. This avoids the issue of violating the proportional hazards assumption. While no subgroup Kaplen-Meier data for intermediate or poor risk groups was identified, comparison of HRs in these subgroups was possible. The results from this alternative method (Appendix D1.1) showed that the results are consistent with the overall population analyses.

In addition, resource use and cost inputs were populated using data reflecting UK clinical practice. Finally, the model concept, inputs, and outputs were reviewed by oncologists actively treating RCC in the UK, thereby ensuring that the model

assumptions were clinically relevant to the UK setting and that a comprehensive array of costs were accounted for.

These were validated to ensure that they were justifiable on the basis of existing data and clinical opinion and were subjected to sensitivity analysis. The model used parametric curve extrapolation for both PFS and OS, based on a parametric curve ITC. The result of the model was impacted by the assumptions around curve extrapolation. In order to take this into consideration, a 20-year time horizon was applied.

Time on treatment for pazopanib was not identified in the published literature. For pazopanib, the median treatment duration and median PFS data of pazopanib were used to estimate the TTD for pazopanib. This might not necessarily reflect clinical practice in the UK.

The results of this analysis demonstrate that improvements in OS and PFS with cabozantinib translate into longer-term gains in LYs and QALYs compared to sunitinib and pazopanib.

B.4 References

1. Ipsen Pharma. *Cabometyx: Summary of Product Characteristics (Draft) (Commercial in Confidence).* 2017.
2. Choueiri TK, Halabi S, Sanford BL, et al. Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Patients With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial. Published online on 14 November 2016. *J Clin Oncol.* 2017;35(6):591-597.
3. Shen C, Kaelin WG, Jr. The VHL/HIF axis in clear cell renal carcinoma. *Semin Cancer Biol.* 2013;23(1):18-25.
4. Zhou L, Liu XD, Sun M, et al. Targeting MET and AXL overcomes resistance to sunitinib therapy in renal cell carcinoma. *Oncogene.* 2016;35(21):2687-2697.
5. Escudier B, Porta C, Schmidinger M, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2016;27(Suppl 5):v58-v68.
6. Lopez-Beltran A, Scarpelli M, Montironi R, Kirkali Z. 2004 WHO classification of the renal tumors of the adults. *Eur Urol.* 2006;49(5):798-805.
7. Cella D. Beyond traditional outcomes: improving quality of life in patients with renal cell carcinoma. *Oncologist.* 2011;16 Suppl 2(Suppl 2):23-31.
8. Gupta K, Miller JD, Li JZ, Russell MW, Charbonneau C. Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): a literature review. *Cancer Treat Rev.* 2008;34(3):193-205.
9. National Institute for Health and Care Excellence. TA178. 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. 26 August 2009; https://www.nice.org.uk/guidance/ta178. Accessed January 2018.
10. Ko JJ, Choueiri TK, Rini BI, et al. First-, second-, third-line therapy for mRCC: benchmarks for trial design from the IMDC. *Br J Cancer.* 2014;110(8):1917-1922.
11. Heng DY, Xie W, Regan MM, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. *Lancet Oncol.* 2013;14(2):141-148.
12. McKay RR, Kroeger N, Xie W, et al. Impact of bone and liver metastases on patients with renal cell carcinoma treated with targeted therapy. *Eur Urol.* 2014;65(3):577-584.
13. Cancer Research UK. Kidney cancer: types and grades. 2017; http://www.cancerresearchuk.org/about-cancer/kidney-cancer/stages-types-grades/types-grades. Accessed January 2018.
14. Kim DY, Wood CG, Karam JA. Treating the two extremes in renal cell carcinoma: management of small renal masses and cytoreductive nephrectomy in metastatic disease. *Am Soc Clin Oncol Educ Book.* 2014:e214-221.
15. Ljungberg B, Hanbury DC, Kuczyk MA, et al. Renal cell carcinoma guideline. *Eur Urol.* 2007;51(6):1502-1510.

16. National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology: Kidney Cancer, version 2.2018.* 2017.
17. National Institute for Health and Care Excellence. Single technology appraisal. Tivozanib for treating renal cell carcinoma [ID591]. Committee papers 2017. https://www.nice.org.uk/guidance/gid-ta10123/documents/committee-papers. Accessed January 2018.
18. National Institute for Health and Care Excellence. Single technology appraisal (STA). Pazopanib (Votrient®) for the first-line treatment of patients with advanced renal cell carcinoma (RCC). Manufacturer submission. 16 April 2010. https://www.nice.org.uk/guidance/ta215/documents/renal-cell-carcinoma-first-line-metastatic-pazopanib-manufacturer-submission-submission2. Accessed January 2018.
19. National Institute for Health and Care Excellence. TA169. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma. 25 March 2009; https://www.nice.org.uk/guidance/ta169. Accessed January 2018.
20. 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):1091-1096.
21. Heng DY, Xie W, Regan MM, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. *J Clin Oncol.* 2009;27(34):5794-5799.
22. Motzer RJ, Bacik J, Schwartz LH, et al. Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma. *J Clin Oncol.* 2004;22(3):454-463.
23. Mekhail TM, Abou-Jawde RM, Boumerhi G, et al. Validation and extension of the Memorial Sloan-Kettering prognostic factors model for survival in patients with previously untreated metastatic renal cell carcinoma. *J Clin Oncol.* 2005;23(4):832-841.
24. Escudier B, Tannir NM, McDermott DF, et al. LBA5CheckMate 214: Efficacy and safety of nivolumab + ipilimumab (N+I) v sunitinib (S) for treatment-naïve advanced or metastatic renal cell carcinoma (mRCC), including IMDC risk and PD-L1 expression subgroups. *Ann Oncol.* 2017;28(suppl_5):mdx440.029-mdx440.029.
25. Rini BI, Atkins MB, Escudier BJ, et al. Tivo-3: A phase 3, randomized, controlled, multi-center, open-label study to compare tivozanib hydrochloride to sorafenib in subjects with refractory advanced renal cell carcinoma (RCC). *J Clin Oncol.* 2017;35(15_suppl):TPS4600-TPS4600.
26. Ko JJ, Xie W, Kroeger N, et al. The International Metastatic Renal Cell Carcinoma Database Consortium model as a prognostic tool in patients with metastatic renal cell carcinoma previously treated with first-line targeted therapy: a population-based study. *Lancet Oncol.* 2015;16(3):293-300.
27. Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med.* 2013;369(8):722-731.
28. Motzer RJ, Hutson TE, Tomczak P, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med.* 2007;356(2):115-124.

29. Sternberg CN, Davis ID, Mardiak J, et al. Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. *J Clin Oncol.* 2010;28(6):1061-1068.
30. Cancer Research UK. Kidney cancer statistics. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer. Accessed January 2018.
31. Office for National Statistics. Cancer Registration Statistics, England, 2015 final. 2017; https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland. Accessed January 2018.
32. National Cancer Registration and Analysis Service. *TNM stage group by CCG by tumor type for 10+3 tumour types, 2015.* 2015.
33. National Institute for Health and Care Excellence. Nivolumab for previously treated advanced renal cell carcinoma. TA417. 23 November 2016; https://www.nice.org.uk/guidance/ta417. Accessed January 2018.
34. National Institute for Health and Care Excellence. Renal cancer overview. https://pathways.nice.org.uk/pathways/renal-cancer. Accessed January 2018.
35. National Institute for Health and Care Excellence. TA215. Pazopanib for the first-line treatment of advanced renal cell carcinoma. Last updated 1 August 2013; https://www.nice.org.uk/guidance/ta215. Accessed January 2018.
36. Powles T, Albiges L, Staehler M, et al. Updated European Association of Urology Guidelines Recommendations for the Treatment of First-line Metastatic Clear Cell Renal Cancer. *Eur Urol.* 2017;In press.
37. Beuselinck B, Oudard S, Rixe O, et al. Negative impact of bone metastasis on outcome in clear-cell renal cell carcinoma treated with sunitinib. *Ann Oncol.* 2011;22(4):794-800.
38. National Institute for Health and Care Excellence. TA463. Cabozantinib for previously treated advanced renal cell carcinoma. 9 August 2017. https://www.nice.org.uk/guidance/ta463/resources/cabozantinib-for-previously-treated-advanced-renal-cell-carcinoma-pdf-82604903877061. Accessed January 2018.
39. Yu H, Zhou YG, Wang YQ. Efficiency and safety of sorafenib in treatment of renal cell carcinoma of Meta-analysis. [Chinese]. *Chinese Pharmaceutical Journal.* 2014;49(5):427-430.
40. Zhuang QY, Liu F, Chen XG, et al. Meta-analysis of anti-vascular endothelial growth factor agents therapy for advanced renal cell carcinoma. [Chinese]. *Chinese Journal of Evidence-Based Medicine.* 2010;10(6):688-692.
41. Sternberg CN, Eisen T, Tomczak P, Strahs AL, Esteves B, Berkenblit A. Tivozanib in patients treatment-naive for metastatic renal cell carcinoma: a subset analysis of the phase III TIVO-1 study. *J Clin Oncol.* 2013;13(Suppl):A4513.
42. Boccardo F, Rubagotti A, Canobbio L, et al. Interleukin-2, interferon-alpha and interleukin-2 plus interferon-alpha in renal cell carcinoma. A randomized phase II trial. *Tumori.* 1998;84(5):534-539.
43. Tomita Y, Naito S, Sassa N. Sunitinib versus sorafenib as first-line therapy followed by sorefenib and sunitinib for patients with metastatic renal cell carcinoma (RCC) with clear cell histology: A multicenter randomized trial,

CROSS-J-RCC. Poster presented at ASCO GU; 2017 Feb 16-18 2017. 2017; Orlando, FL.
44. Exelixis Inc. *Clinical Study Report: A031203 Randomized Phase II Study Comparing Cabozantinib (NSC #761968 and IND #116059) with Commercially Supplied Sunitinib in Subjects with Previously Untreated Locally Advanced or Metastatic Renal Cell Carcinoma. Report dated 31 July 2017. (Commercial in Confidence).* 2017.
45. Choueiri TK, Halabi S, Sanford B, et al. PR CABOzantinib versus SUNitinib (CABOSUN) as initial targeted therapy for patients with metastatic renal cell carcinoma (mRCC) of poor and intermediate risk groups: Results from ALLIANCE A031203 trial. *Ann Oncol.* 2016;27(Suppl 6).
46. Chopra M. Annual Congress of the European Society for Medical Oncology (ESMO): Copenhagen, Denmark; 7-11 October 2016. *Target Oncol.* 2016;11(6):705-709.
47. Choueiri T, Hessel C, Halbi S, et al. Progression-free survival by independent review and overall survival update for the Alliance A031203 CABOSUN trial of cabozantinib vs sunitinib in metastatic renal cell carcinoma. European Society for Medical Oncology Congress; September, 2017; Madrid, Spain.
48. Food and Drug Administration. *Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics.* 2007.
49. Guo J, Jin J, Huang Y, et al. Comparison of PFS and safety for Asian compared to North American and European populations in the phase III trial of pazopanib versus sunitinib in patients with treatment-naive RCC (COMPARZ). *J Clin Oncol.* 2013;31(6 Suppl1):366.
50. Motzer R, Hutson TE, Reeves J, et al. Randomized, open-label, phase III trial of pazopanib versus sunitinib in first-line treatment of patients with Metastatic Renal Cell Carcinoma (MRCC): Results of the COMPARZ trial. *Ann Oncol.* 2012;23(Suppl 9):ix13.
51. Motzer RJ, Hutson TE, McCann L, Deen K, Choueiri TK. Overall survival in renal-cell carcinoma with pazopanib versus sunitinib. *N Engl J Med.* 2014;370(18):1769-1770.
52. Novartis Pharmaceuticals. Pazopanib versus sunitinib in the treatment of locally advanced and/or metastatic renal cell carcinoma (COMPARZ). *ClinicalTrials gov.* 2017. https://ClinicalTrials.gov/show/NCT00720941. Accessed January 2018.
53. Motzer RJ, Hutson TE, Tomczak P, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. *J Clin Oncol.* 2009;27(22):3584-3590.
54. Tomita Y, Naito S, Sassa N, et al. Sunitinib versus sorafenib as first-line therapy for patients with metastatic renal cell carcinoma with favorable or intermediate MSKCC risk factors: A multicenter randomized trial, CROSS-J-RCC. *J Clin Oncol.* 2014;32(4 Suppl1):502.
55. Sternberg CN, Hawkins RE, Wagstaff J, et al. A randomised, double-blind phase III study of pazopanib in patients with advanced and/or metastatic renal cell carcinoma: final overall survival results and safety update. *Eur J Cancer.* 2013;49(6):1287-1296.
56. Sternberg CN, Hawkins RE, Szczylik C, et al. Randomized, double-blind phase III study of pazopanib in patients with advanced/metastatic renal cell

carcinoma (mRCC): Final overall survival (OS) results. *Ann Oncol.* 2010;21:viii10.
57. Escudier B, Szczylik C, Hutson TE, et al. Randomized phase II trial of first-line treatment with sorafenib versus interferon alfa-2a in patients with metastatic renal cell carcinoma. *J Clin Oncol.* 2009;27(8):1280-1289.
58. Negrier S, Jager E, Porta C, et al. Efficacy and safety of sorafenib in patients with advanced renal cell carcinoma with and without prior cytokine therapy, a subanalysis of TARGET. *Med Oncol.* 2010;27(3):899-906.
59. Escudier B, Pluzanska A, Koralewski P, et al. Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. *Lancet.* 2007;370(9605):2103-2111.
60. Escudier B, Bellmunt J, Negrier S, et al. Phase III trial of bevacizumab plus interferon alfa-2a in patients with metastatic renal cell carcinoma (AVOREN): Final analysis of overall survival. *J Clin Oncol.* 2010;28(13):2144-2150.
61. Rini BI, Halabi S, Rosenberg JE, et al. Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206. *J Clin Oncol.* 2008;26(33):5422-5428.
62. Rini BI, Halabi S, Rosenberg JE, et al. Phase III trial of bevacizumab plus interferon alfa versus interferon alfa monotherapy in patients with metastatic renal cell carcinoma: Final results of CALGB 90206. *J Clin Oncol.* 2010;28(13):2137-2143.
63. Motzer RJ, Nosov D, Eisen T, et al. Tivozanib versus sorafenib as initial targeted therapy for patients with metastatic renal cell carcinoma: results from a phase III trial. *J Clin Oncol.* 2013;31(30):3791-3799.
64. Negrier S, Gravis G, Perol D, et al. Temsirolimus and bevacizumab, or sunitinib, or interferon alfa and bevacizumab for patients with advanced renal cell carcinoma (TORAVA): A randomised phase 2 trial. *Lancet Oncol.* 2011;12(7):673-680.
65. Hudes G, Carducci M, Tomczak P, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. *N Engl J Med.* 2007;356(22):2271-2281.
66. Eichelberg C, Vervenne WL, De SM, et al. SWITCH: A Randomised, Sequential, Open-label Study to Evaluate the Efficacy and Safety of Sorafenib-sunitinib Versus Sunitinib-sorafenib in the Treatment of Metastatic Renal Cell Cancer. *Eur Urol.* 2015;68(5):837-847.
67. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-560.
68. National Institute for Health and Care Excellence. Process and methods [PMG9]. Guide to the methods of technology appraisal 2013. *National Institute for Health and Care Excellence.* 2013. http://nice.org.uk/process/pmg9. Accessed January 2018.
69. Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med.* 2015;373(19):1814-1823.
70. Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. *Lancet Oncol.* 2016;17(7):917-927.
71. Stukalin I, Alimohamed N, Heng DY. Contemporary Treatment of Metastatic Renal Cell Carcinoma. *Oncol Rev.* 2016;10(1):295.
72. Rini BI, Vogelzang NJ. Future Challenges for Drug Development in Renal Cell Carcinoma. *J Clin Oncol.* 2017;35(6):577-579.

73. Choueiri TK, Halabi S, Morris MJ, George D. Reply to B. Rini et al and S. Buti et al. *J Clin Oncol.* 2017;35(16):1859-1860.
74. Motzer RJ, Barrios CH, Kim TM, et al. Phase II randomized trial comparing sequential first-line everolimus and second-line sunitinib versus first-line sunitinib and second-line everolimus in patients with metastatic renal cell carcinoma. *J Clin Oncol.* 2014;32(25):2765-2772.
75. Rini BI, Hutson TE, Figlin RA, et al. Sunitinib in patients with metastatic renal cell carcinoma: Clinical outcome according to IMDC risk group. *J Clin Oncol.* 2017;35(15,suppl):abstract 4584.
76. Schnadig ID, Hutson TE, Chung H, et al. Dosing patterns, toxicity, and outcomes in patients treated with first-line sunitinib for advanced renal cell carcinoma in community-based practices. *Clin Genitourin Cancer.* 2014;12(6):413-421.
77. Amdahl J, Diaz J, Sharma A, Park J, Chandiwana D, Delea TE. Cost-effectiveness of pazopanib versus sunitinib for metastatic renal cell carcinoma in the United Kingdom. *PLoS One.* 2017;12(6):e0175920.
78. Hoyle M, Green C, Thompson-Coon J, et al. Cost-effectiveness of temsirolimus for first line treatment of advanced renal cell carcinoma. *Value Health.* 2010;13(1):61-68.
79. All Wales Medicines Strategy Group (AWMSG). *AWMSG advice superseded by NICE guidance (TA169). Final Appraisal Report. Sunitinib (Sutent®).* 2007.
80. Kilonzo M, Hislop J, Elders A, et al. Pazopanib for the first-line treatment of patients with advanced and/or metastatic renal cell carcinoma: A NICE single technology appraisal. *Pharmacoeconomics.* 2013;31(1):15-24.
81. Thompson Coon J, Hoyle M, Green C, 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.
82. National Institute for Health and Care Excellence. NICE technology appraisal guidance [GID-TA10123]. Tivozanib for treating renal cell carcinoma [ID591]. *National Institute for Health and Care Excellence.* 2017. https://www.nice.org.uk/Guidance/gid-ta10123. Accessed January 2018.
83. National Institute for Health and Care Excellence. NICE technology appraisal guidance [TA333]. Axitinib for treating advanced renal cell carcinoma after failure of prior systemic treatment. *National Institute for Health and Care Excellence.* 2015. https://www.nice.org.uk/guidance/ta333. Accessed January 2018.
84. National Institute for Health and Care Excellence. NICE technology appraisal guidance [TA219]. Everolimus for the second-line treatment of advanced renal cell carcinoma. *National Institute for Health and Care Excellence.* 2011. https://www.nice.org.uk/guidance/ta219. Accessed 20/11/2017.
85. GlaxoSmithKline. Clinical study report. Study VEG108844, a study of pazopanib versus sunitinib in the treatment of subjects with locally advanced and/or metastatic renal cell carcinoma. *GlaxoSmithKline.* 2013. https://www.gsk-clinicalstudyregister.com/study/108844#csr. Accessed January 2018.
86. Sternberg CN, Hawkins RE, Szczylik C, et al. A randomized, double-blind phase III study (VEG105192) of pazopanib (paz) versus placebo (pbo) in patients with advanced/metastatic renal cell carcinoma (mRCC): Updated safety results. *J Clin Oncol.* 2011;29(7 Suppl1):313.

87. Remak E, Charbonneau C, Negrier S, Kim ST, Motzer RJ. Economic evaluation of sunitinib malate for the first-line treatment of metastatic renal cell carcinoma. *J Clin Oncol.* 2008;26(24):3995-4000.
88. Parasuraman S, Hudes G, Levy D, Strahs A, Moore L, DeMarinis R. Comparison of quality-adjusted survival in patients with advanced renal cell carcinoma receiving first-line treatment with temsirolimus (TEMSR) or interferon- (IFN) or the combination of IFN+TEMSR. *J Clin Oncol.* 2008;25(18Suppl):5049.
89. European Medicines Agency. EPAR summary for the public. Sutent. *European Medicines Agency.* 2014. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/000687/WC500057689.pdf. Accessed January 2018.
90. European Medicines Agency. EPAR summary for the public. Votrient. *European Medicines Agency.* 2012. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/001141/WC500094273.pdf. Accessed January 2018.
91. Ouwens MJ, Philips Z, Jansen JP. Network meta-analysis of parametric survival curves. *Res Synth Methods.* 2010;1(3-4):258-271.
92. Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Med Res Methodol.* 2011;11:61.
93. Latimer N. NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. *NICE Decision Support Unit.* 2013. http://www.nicedsu.org.uk. Accessed 8/24/2016.
94. The GlaxoSmithKline group of companies. *Study VEG108844, a study of pazopanib versus sunitinib in the treatment of subjects with locally advanced and/or metastatic renal cell carcinoma: overall survival update. Report dated 13 Mar 2014. (GlaxoSmithKline document number 2013N180282_00).*
95. Hernandez Alava M, Wailoo A, Pudney S. Methods for mapping between the EQ-5D-5L and the 3L for technology appraisal. *NICE Decision Support Unit.* 2017. http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Mapping-5L-to-3L-DSU-report.pdf. Accessed January 2018.
96. Castellano D, del Muro XG, Perez-Gracia JL, et al. Patient-reported outcomes in a phase III, randomized study of sunitinib versus interferon-{alpha} as first-line systemic therapy for patients with metastatic renal cell carcinoma in a European population. *Ann Oncol.* 2009;20(11):1803-1812.
97. Lai JS, Beaumont JL, Diaz J, Khan S, Cella D. Validation of a short questionnaire to measure symptoms and functional limitations associated with hand-foot syndrome and mucositis in patients with metastatic renal cell carcinoma. *Cancer.* 2016;122(2):287-295.
98. Amdahl J, Diaz J, Park J, Nakhaipour HR, Delea TE. Cost-effectiveness of pazopanib compared with sunitinib in metastatic renal cell carcinoma in Canada. *Curr Oncol.* 2016;23(4):e340-354.
99. Goebell PJ, Münch A, Müller L, et al. A cross-sectional investigation of fatigue in advanced renal cell carcinoma treatment: results from the FAMOUS study. *Urol Oncol.* 2014;32(3):362-370.

100. Exelixis. *Clinical study report. XL184-308. A phase 3, randomized, controlled study of cabozantinib (XL184) vs everolimus in subjects with metastatic renal cell carcinoma that has progressed after prior VEGFR tyrosine kinase inhibitor therapy. (Commercial in Confidence).* 12/11/2015 2015.
101. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health.* 2010;13(5):509-518.
102. National Health Service. *BNF, NHS indicative price.* 2017 2017.
103. NHS Improvement. NHS Reference Costs 2016/2017. 2017. https://improvement.nhs.uk/resources/reference-costs/. Accessed January 2018.
104. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. *Nuffield Trust.* 2014. https://www.nuffieldtrust.org.uk/research/exploring-the-cost-of-care-at-the-end-of-life. Accessed January 2018.
105. Curtis L, Burns A. Unit costs of health and social care 2016. *Personal Social Services Research Unit.* 2016. http://www.pssru.ac.uk/pub/uc/uc2016/full.pdf. Accessed January 2018.
106. National Institute for Health and Care Excellence. Clinical guideline [CG144]. Venous thromboembolic diseases: diagnosis, management and thrombophilia testing. *National Institute for Health and Care Excellence.* 2012. https://www.nice.org.uk/guidance/cg144. Accessed January 2018.
107. Inflation.eu. CPI inflation Great Britain. *Inflation eu.* 2016. www.inflation.eu. Accessed January 2018.
108. Rini BI, Escudier B, Tomczak P, et al. Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. *Lancet.* 2011;378(9807):1931-1939.
109. National Institute for Health and Care Excellence. Single technology appraisal. Nivolumab for treated or metastatic renal cell carcinoma [ID853]. Committee Papers. https://www.nice.org.uk/guidance/ta417/documents/committee-papers-4. Accessed January 2018.

B.5 Appendices

Appendices D to L are presented in a consolidated document (the Appendices document). The SmPC is supplied as a separate document.

Appendix	Location
Appendix C – SmPC and EPAR	The SmPC has been submitted as a separate document, as part of the reference pack. The EPAR is not yet available.
Appendix D1.1 – Systematic literature review, indirect treatment comparison (ITC) methodology and supplementary ITC results Appendix D1.2 – Participant flow in the CABOSUN trial Appendix D1.3 – Quality assessment of the CABOSUN trial	Appendices document
Appendix E – Subgroup analyses	
Appendix F – Adverse reactions	
Appendix G – Published cost-effectiveness studies	
Appendix H – Health-related quality of life studies	
Appendix I – Cost and healthcare resource identification	
Appendix J – Clinical outcomes and disaggregated results	
Appendix K – Checklist of confidential information	
Appendix L – Additional data from the CABOSUN trial	

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Single technology appraisal
Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma [ID1208]

Dear [redacted]

The Evidence Review Group, Southampton Health Technology Assessments Centre and the technical team at NICE have looked at the submission received on 19th February from Ipsen Ltd UK. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **8th March**. Your response and any supporting documents should be uploaded to NICE Docs/Appraisals https://appraisals.nice.org.uk/request/45475

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Alan Lamb, Technical Lead (alan.lamb@nice.org.uk). Any procedural questions should be addressed to Jeremy Powell, Project Manager (jeremy.powell@nice.org.uk).

Yours sincerely

Elisabeth George
Associate Director – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

A1. The systematic review search sources do not appear to have included conferences, although some abstracts were retrieved. Please clarify whether a systematic search was conducted for conference abstracts.

A2. The company submission states that advanced RCC includes both locally advanced RCC that cannot be removed by surgery, and metastatic RCC. In the CABOSUN trial all patients had metastatic disease. Therefore, we assume that there were no patients with locally advanced RCC enrolled in the trial. Please can you confirm that this is correct?

A3. **Priority question.** We note from Figure 5 of the company submission that the Kaplan–Meier curves for PFS in the CABOSUN trial visually appear parallel. Appendix D Table 12 reports conclusions about proportional hazards assumptions, with the assumption not appearing to hold for PFS. Please can you provide more detail on the assessments made, and supply the Schoenefeld plots and log cumulative plots for OS and PFS in both the CABOSUN and COMPARZ trials. Please also provide the same information for TTD in the CABOSUN trial.

A4. Table 9 of the company submission: Please explain how the Oncology Patient Enrollment Network (OPEN) registration system works:

a. How does it ensure allocation of patients is truly random?

b. Is the allocation sequence concealed until patients are assigned to treatment? If so, how?

c. We also note that the dynamic allocation method was used as part of randomisation (Appendix D1.3, Table 15). Please can you describe this method and how it was used in the trial?

A5. Table 9 of the company submission: Please explain the process that was used to ensure that the Independent Radiology Committee (IRC) was blinded to the treatment assignment corresponding to the radiographic images.

A6. Missing data for ORR: Table 11 states that no values were imputed for patients for whom baseline and post-baseline radiographic images were not available. Does this mean that these data were censored at baseline? (They are listed as censored in Table 12).

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A7. According to Table 12 of the company submission, there was an imbalance between the cabozantinib and sunitinib arms in (i) the proportions of patients who had ≥2 missed "adequate tumour assessments" before a PFS event, and (ii) the proportions who had no post-baseline "adequate tumour assessments". What are the reasons for these differences?

A8. Section B.2.6 of the company submission states that, for the IRC assessment of PFS, 13 patients did not have complete data for radiographic images or tumour assessments. Figure 8, showing the percentage change in tumour size, states that 24 patients did not have adequate post-baseline imaging assessments. Please explain the difference in the number of missing observations between these two outcomes.

A9. Subgroup analyses.

a. Was any adjustment made for multiple testing among the 16 subgroup analyses reported in Figure 53 in the Appendix?

b. Please supply the subgroup analyses results for OS.

A10. In relation to OS and PFS, it is stated that the plausibility of different extrapolations was assessed by visual inspection by oncologists currently practising within the NHS in England. Please can you elaborate on this process, including the number of oncologists consulted, their clinical speciality, etc.

A11. On page 56 of the company submission it is stated that there was "use of two independent reviewers in the SLR". Please can you clarify which of these processes this applied to: screening titles and abstracts, screening full texts, and/or data extraction and critical appraisal of included studies?

A12. On page 56 of the company submission it is stated "Additionally, the SLR may have been impacted by the decision to exclude certain studies due to data availability". We assume that this is referring to the 9 full text articles listed in Figure 3 excluded as "article not obtained". Further, Table 8 in the appendix lists 9 references which appear to have been excluded as not obtainable. We assume that these are the same 9 listed in Figure 3. Please can you confirm this, and whether there are any other studies that were excluded due to data availability, and also what the impact may have been of any exclusions?

A13. Please can you confirm whether you are aware of any planned/ongoing trials or studies of cabozantinib for this indication?

A14. What is included in the 'other' reason for discontinuation from study treatment in Figure 52 in appendix D1.2?

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A15. The critical appraisal of study quality (Table 15, appendix D1.3) states that there were no unexpected imbalances in dropouts in the CABOSUN trial. However, Figure 52 in appendix D1.2 shows 'withdrawal by subject during treatment' as more than double in the sunitinib arm (n=7) compared with the cabozantinib arm (n=3). While these numbers are low, please explain why you did not consider this an imbalance.

A16. Table 26 lists solicited adverse events, and unsolicited related AEs reported by ≥10% of patients in either treatment arm. The text on page 59 lists grade 3 or 4 AEs occurring in ≥5% of patients. Some of the AEs in the text as occurring in ≥5% patients (e.g. stomatitis in the cabozantinib arm) are listed in the table as reported by ≥10% of patients in either treatment arm. Please can you explain the criteria for inclusion of adverse events on page 59 and Table 26.

Subsequent anticancer therapies

A17. On page 44 of the company submission, it is stated that subsequent non-radiation anticancer therapy was used by 57% and 58% of patients in the cabozantinib and sunitinib arms, respectively. Table 16 states that first systemic subsequent anticancer therapy was used in 37 (46.8%) patients in the cabozantinib arm. We presume that the figure of 57% stated on page 44 includes first and subsequent treatment lines, and/or the figures in Table 50 in Appendix L are from a different data-cut – please can you confirm?

A18. In Table 56 of the company submission, the proportion of patients who received nivolumab as a subsequent treatment was 13% and 15% in cabozantinib and sunitinib patients, respectively. In Appendix L (Table 50) the proportion of patients receiving nivolumab as subsequent treatment is 5% and 8% in the cabozantinib and sunitinib arms, respectively. We assume that in Table 56 of the company submission you have summed the percentage of patients receiving nivolumab with the percentage of patients who received PD-1 inhibitor from Appendix Table 50. Please can you confirm if our assumption is correct?

A19. Please can you report the number (%) of patients who received each subsequent line of systemic anticancer therapy (e.g. second, third, fourth etc)?

A20. In the clinical study report, it is stated that "The SAP, available in Appendix 16.1.9.1, describes the statistical analyses that were performed by Exelixis". This appendix is not available to the ERG. Please can you supply the SAP (statistical analysis plan)?

Indirect treatment comparison (ITC)

A21. For the Ouwens et al model and the fractional polynomials model please provide more information about the Bayesian methods used including the prior probability distributions (e.g. vague, informative, non-informative, the rationale for their choice, and any sensitivity analyses on these), the likelihood distribution, the number of

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

iterations used for burn in and inferences, and the methods for assessing convergence.

A22. **Priority question**. Please clarify the interval used for dividing up the follow-up period in the fractional polynomial model. Please supply the tabulated hazard ratios and 95% credible intervals for each fractional polynomial model for each interval time point. Please also supply hazard ratio plots for each fractional polynomial model with credible intervals to allow visual inspection.

A23. **Priority question**. Please provide fractional polynomial results (in terms of curves, and hazard ratios and 95% credible intervals for each time point) based on a random-effect model.

A24. **Priority question**. Please explain why you chose to test fractional polynomials with a relatively narrow range of powers (P1 and P2 in the range -1 to +1). Given that none of the overall survival curves appears to reflect the shape of the CABOSUN Kaplan–Meier curves shown in Figure 6, did you consider trying other functional forms?

A25. Page 89 of the company submission states that "The FP 1st order models have lower DIC than the FP 2nd models, and therefore were not included in the scenario analyses". We note from Table 24 that the fractional polynomials 1st order models have *higher* DIC values than the 2nd order models. Is this a typographical error?

A26. **Priority question.** We are interested to know whether the results of the ITC for the comparison of cabozantinib with sunitinib and pazopanib would differ if the wider primary evidence network shown in Figure 9 was used, rather than the restricted network shown in Figure 11. Please provide full results (random and fixed effect) of the ITC for the primary evidence network using the Ouwens et al and fractional polynomial methods, accompanied by an assessment of statistical and clinical heterogeneity, inconsistency, and methodological study quality.

A27. **Priority question**. For overall and progression-free survival, please provide graphs showing the CABOSUN and regenerated COMPARZ Kaplan–Meier graphs overlaid on the fitted survival curves (in appendix D) for each of the ITC parametric and fractional polynomial models.

Section B: Clarification on cost-effectiveness data

B1. **Priority question**. Please provide Kaplan–Meier data from CABOSUN and the regenerated Kaplan–Meier data from COMPARZ that you used to fit the parametric and fractional polynomial overall and progression-free survival models.

B2. Please justify your choice of exponential distribution for the direct comparison of overall survival between cabozantinib and sunitinib in the base case. This is

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

inconsistent with the conclusion that the proportional hazards assumption does not hold for the CABOSUN OS data. The use of separate exponential curves also appears to give a poor fit for the Kaplan–Meier data in Figure 13.

B3. **Priority question**. Please can you clarify your assumptions regarding time to treatment discontinuation for pazopanib in the model. On page 142 (B.3.11) it is stated that; "Time on treatment for pazopanib was not identified in the published literature. For pazopanib, the median treatment duration and median PFS data of pazopanib were used to estimate the TTD for pazopanib". We assume this is a typographical error and that the median treatment duration and PFS data of sunitinib were used as an estimate, please can you confirm this assumption? This assumption does not seem to match the formulae in the model (column S in 'E.Pazo.ITC' and 'E.Pazo.FP).

B4. The utility values reported in Table 43 of the company submission for progression-free (0.78) and progressed (0.705) states are different from those used in the model: 0.77 for progression-free and 0.72 for progressed. Please clarify these differences.

Section C: Textual clarifications and additional points

C1. Table 15: Please clarify the meaning of footnote d, which appears ambiguous.

C2. Table 13 footnote states "HR <1 indicates OS in favour of sunitinib". We presume this should say "HR >1" or "… in favour of cabozanitinib"?

Ipsen Ltd response to clarification questions – 8 March 2018

Single technology Appraisal
Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma [ID1208]

Please note that, in responding to the clarification questions, we have provided as much information as possible and appreciate that, due to the data requested, the response document is quite lengthy. In the event further information / clarification is required please do not hesitate to contact us.

Section A: Clarification on effectiveness data

A1.

The systematic review search sources do not appear to have included conferences, although some abstracts were retrieved. Please clarify whether a systematic search was conducted for conference abstracts.

Response: A systematic search for conference abstracts was not conducted. Conference abstracts were retrieved through the search in EMBASE and Cochrane databases, and via the systematic review of the pazopanib manufacturer submission document.

A2.

The company submission states that advanced RCC includes both locally advanced RCC that cannot be removed by surgery, and metastatic RCC. In the CABOSUN trial all patients had metastatic disease. Therefore, we assume that there were no patients with locally advanced RCC enrolled in the trial. Please can you confirm that this is correct?

Response: We can confirm that all RCC patients in the CABOSUN trial had metastatic disease.

The inclusion criteria concerning stage of disease in the CABOSUN protocol stated that patients needed to have locally advanced (defined as disease not amenable to curative surgery or radiation therapy) or metastatic RCC (equivalent to stage IV RCC, according to AJCC staging). It is important to note that this inclusion criterion is consistent with other trials in this disease area. Each trial will have varying proportions of patients with inoperable locally advanced or metastatic RCC, but the proportion of patients with inoperable locally advanced disease is significantly less than those with metastatic disease. Despite this, these trials are accepted by regulatory authorities and the clinical community to meet the benefit-risk criteria to be used in advanced RCC patients (those with either inoperable locally advanced or metastatic RCC).

A3.

Priority question. We note from Figure 5 of the company submission that the Kaplan–Meier curves for PFS in the CABOSUN trial visually appear parallel. Appendix D Table 12 reports conclusions about proportional hazards assumptions, with the assumption not appearing to hold for PFS. Please can you provide more detail on the assessments made, and supply the Schoenfeld plots and log cumulative plots for OS and PFS in both the CABOSUN and COMPARZ trials. Please also provide the same information for TTD in the CABOSUN trial.

Response: The proportional hazards assumption for the Kaplan-Meier curves for PFS in the CABOSUN trial was checked using statistical tests and graphical diagnostics based on the scaled Schoenfeld residuals (Figure 1), as well as cumulative hazard plots (Figure 2). The proportionality test shows that the proportional hazard assumption is not violated (Text Box 1). However, when we look at scaled Schoenfeld plots, we find an increasing trend followed by a decreasing trend. Also, the log-cumulative plot shows roughly parallel curves and, therefore, we think the separate fitting is more appropriate, as it doesn't require the assumption of proportionality. The scaled Schoenfeld residuals and log-cumulative plots are shown from Figure 1 to Figure 10.

Text Box 1. CABOSUN PFS proportionality test

```
ph.assumption(PFS,  1, 0)
```

```
##  [1]  "Scaled  Schoenfeld:  p<0.05  =  PH  assumption  not  hold"
##                                  rho    chisq                    p
## trt 0.0446  0.18 0.672
```

Figure 1. CABOSUN PFS graphs – scaled Schoenfeld

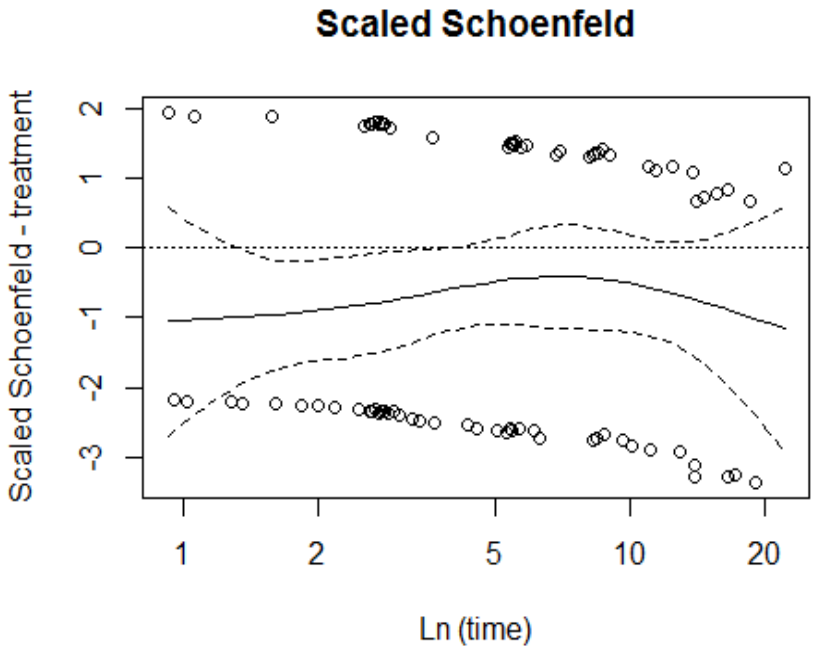

Scaled Schoenfeld
Scaled Schoenfeld - treatment
2
1
0
-1
-2
-3
1
2
5
10
20
Ln (time)

Figure 2. CABOSUN PFS graphs – Log-cumulative plot

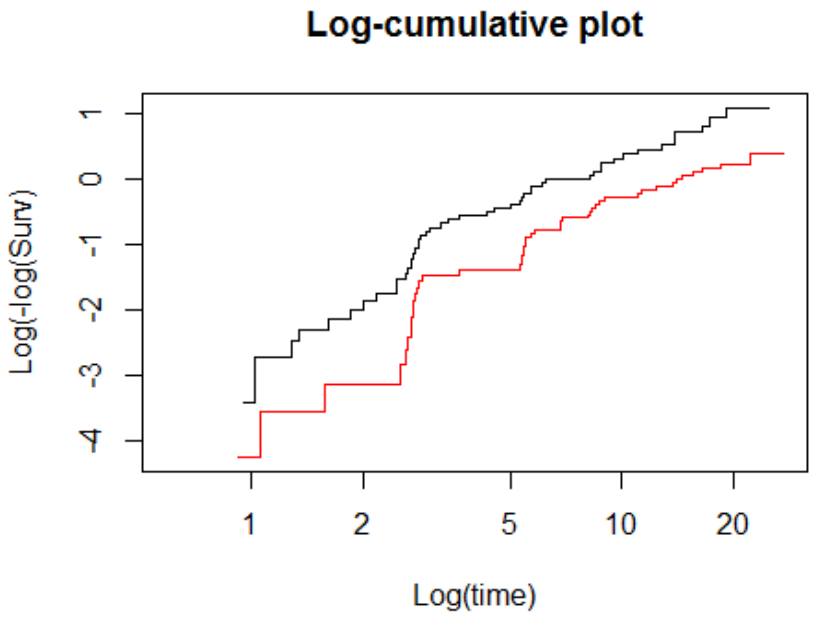


Figure 3. CABOSUN OS graphs – scaled Schoenfeld

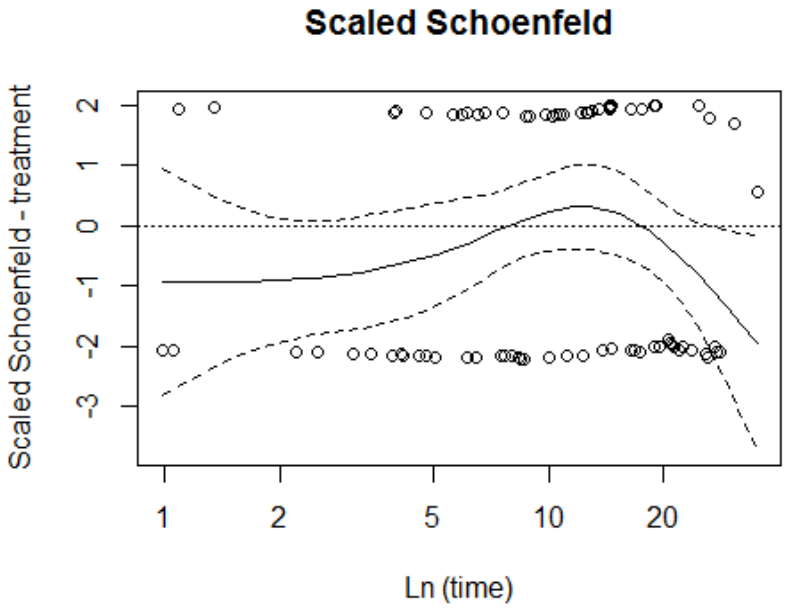


Figure 4. CABOSUN OS graphs – Log-cumulative plot

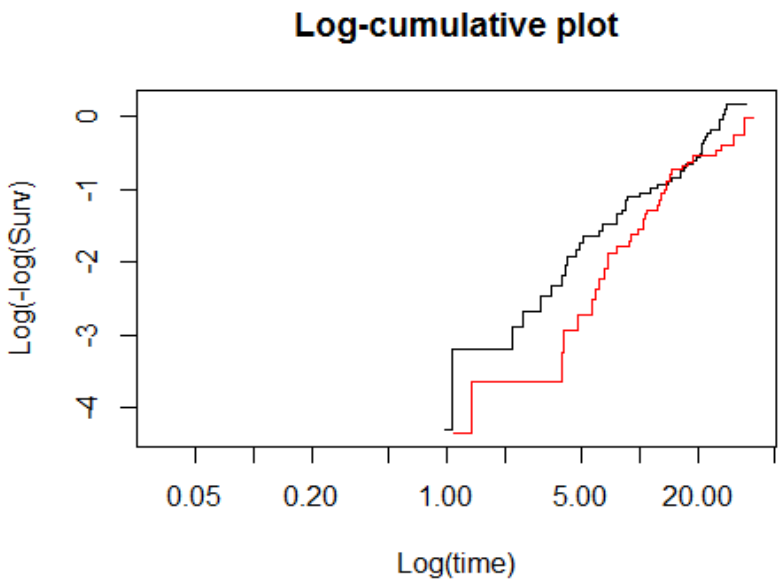


Figure 5. CABOSUN TTD graphs – scaled Schoenfeld

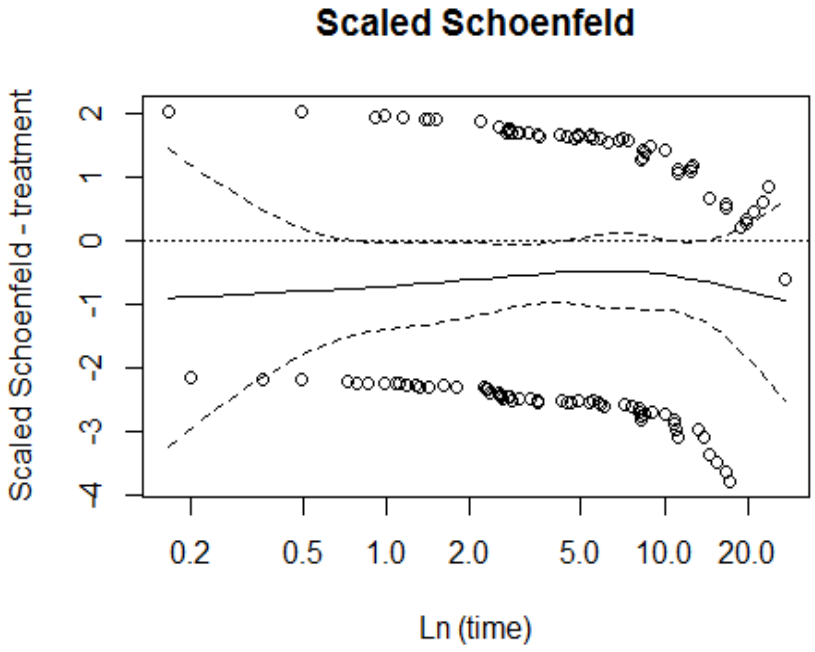


Figure 6. CABOSUN TTD graphs – Log-cumulative plot

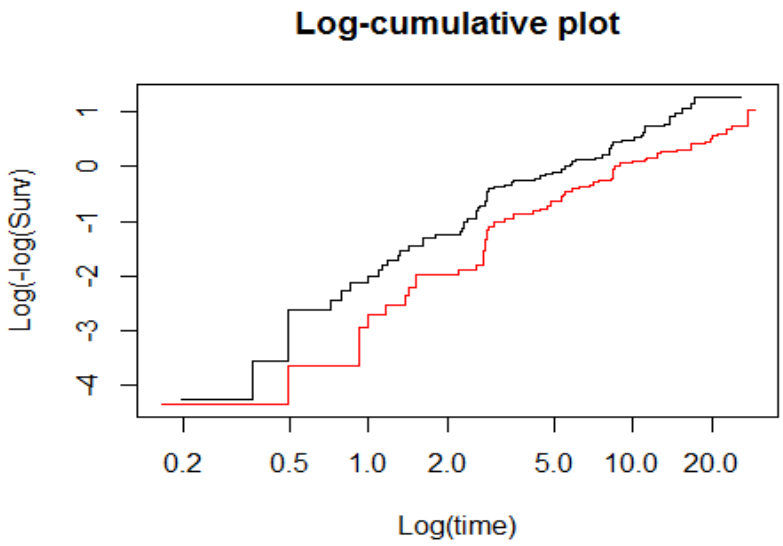


Figure 7. COMPARZ PFS graphs – scaled Schoenfeld

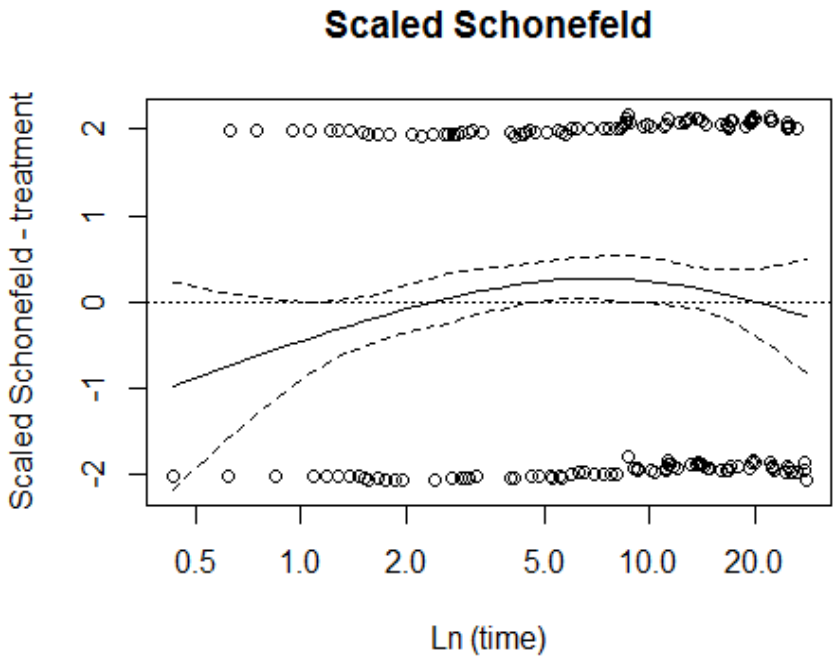


Figure 8. COMPARZ PFS graphs – Log-cumulative plot

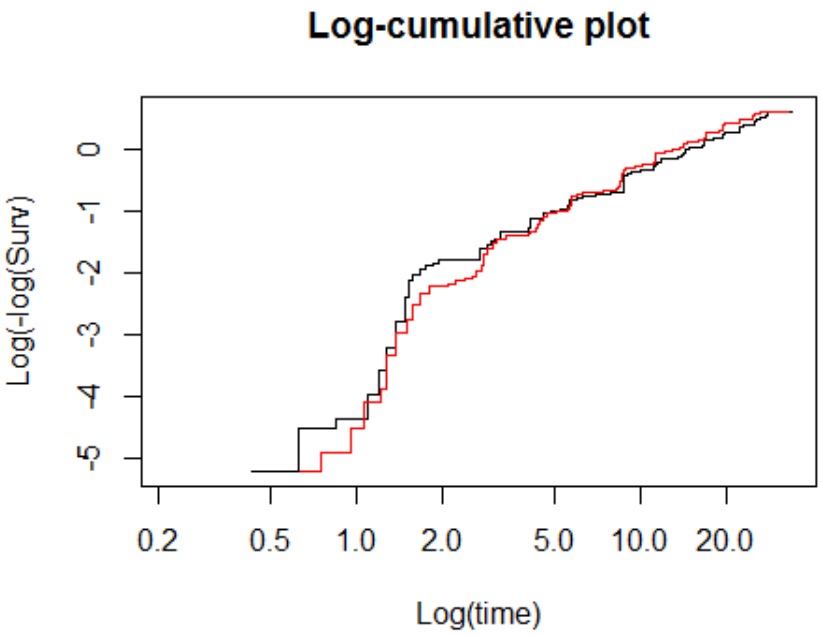


Figure 9. COMPARZ OS graphs – scaled Schoenfeld

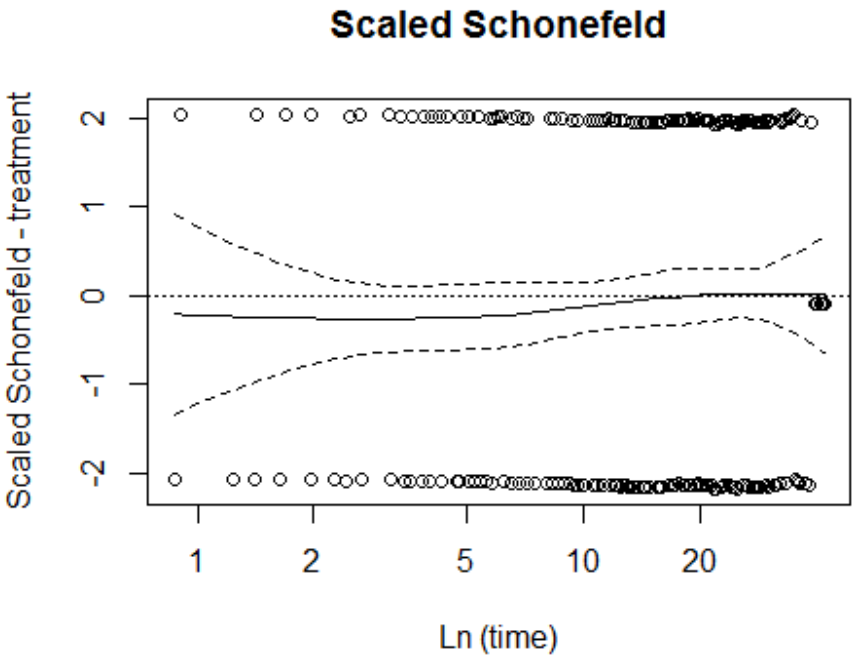
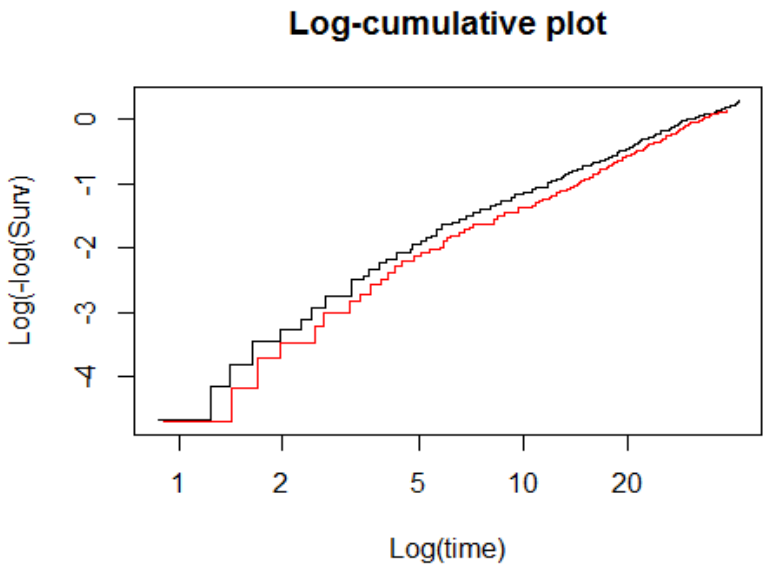


Figure 10. COMPARZ OS graphs – Log-cumulative plot



A4.

Table 9 of the company submission: Please explain how the Oncology Patient Enrollment Network (OPEN) registration system works:

a. How does it ensure allocation of patients is truly random?

b. Is the allocation sequence concealed until patients are assigned to treatment? If so, how?

c. We also note that the dynamic allocation method was used as part of randomisation (Appendix D1.3, Table 15). Please can you describe this method and how it was used in the trial?

Responses:

a) CABOSUN was conducted by the Alliance for Clinical Trials in Oncology under the National Cancer Institute-Cancer Therapy Evaluation Programme (NCI-CTEP) US Investigational New Drug (IND) application. Subject enrollment used the web-based Oncology Patient Enrollment Network (OPEN) registration system to randomise patients. OPEN is the system used to enrol patients into NCI-sponsored Network Group clinical trials. It allows study sites to enrol patients on a 24/7 basis and provides each site with a confirmation of registration (randomisation) and treatment information.

To register a patient, the investigator was required to access the system and enter details on the patient's eligibility as per the Clinical Report Form.

Further details on OPEN can be found at: https://open.ctsu.org

b) The allocation sequence was concealed until the time the site entered the patient's details onto the system and received back details on the treatment arm to which the patient was assigned.

At the time of enrolment, subjects were stratified as follows:

- Bone metastases at baseline
    - Yes
    - No
- IMDC risk categories:
    - Intermediate (1-2 risk factors)
    - Poor (3 or more risk factors)

c) Dynamic allocation (DA) procedures balance treatment arms across baseline prognostic factors. In CABOSUN random assignment was stratified by IMDC risk category (intermediate or poor) and presence of bone metastases (yes or no).

A5.

Table 9 of the company submission: Please explain the process that was used to ensure that the Independent Radiology Committee (IRC) was blinded to the treatment assignment corresponding to the radiographic images.

Response: The IRC reviewed all available radiographic studies and was blinded to treatment identity and to clinical data that could have led to inadvertent unblinding.

MedQIA was contracted to perform image data collection, quality control review and storage, and also to provide an independent assessment of radiographic time-point response for all subjects.

Each site was required to send all images to MedQIA who, prior to release to the independent reviewers, removed all traces of identification (patient details, assessment date, number, etc) from each image according to HIPAA, IRB or Ethics Committee (EC) guidelines, GCP and local regulatory requirements. A QC step was subsequently performed to ensure appropriate de-identification and digitization.

A standard dual reader oncology read paradigm, using validated software tools, was used. At least two independent primary readers were required to perform the work with an additional independent reader identified should one of the primary readers became unavailable.

All readers were blinded to the following:

- Clinical Site image interpretations, including choice of radiographic target and non-targeted lesions and the identification of radiographic new lesions, and response assessment
- Clinical data
- Read results from other independent central reviewers
- Subject demographics
- Treatment arm
- The number of time points for a subject (to eliminate progression bias by the reader). Readers were only presented with the next available time point after

signing off the prior time point, and therefore could not determine in advance whether the imaging time point is an ongoing or final imaging time point for a subject.

An overview of the Read process is as follows:

- The reader was required to log onto the MedQIA database using an identifier code and unique password.
- The reader was presented with the full image data set of the sequences being analysed with the lesions highlighted. After completing the review of the image, its lesion segmentations and CRF, the reader had the option to make further edits, or save and approve their annotations. Once case is approved, the annotations could not be modified without audit trail.
- Following Reader approval of annotations, time point response assessment (TPR) was computed automatically using MedQIA validated software.

Please note that there is an error in the CSR page 53, Section 9.4.7 which, in relation to the IRC review states: '*Treatment assigned and treatment received were excluded from the data review output unless necessary to conduct the review*'. This text was inserted in error. There were no circumstances in CABOSUN where the IRC received unblinded treatment information.

A6.

Missing data for ORR: Table 11 states that no values were imputed for patients for whom baseline and post-baseline radiographic images were not available. Does this mean that these data were censored at baseline? (They are listed as censored in Table 12).

Response: We have assumed that, as Table 12 is for PFS by IRC, the above question should read 'Missing data for PFS'.

In the retrospective IRC assessment of PFS (and ORR), no values were imputed for patients for whom a complete set of baseline and post-baseline radiographic images were not available.

The following FDA censoring rules were used in analyses of PFS:

- Only adequate tumour assessments (ATAs) with an overall response of CR, PR, SD or PD were considered in the determination of progression/censoring dates.

- Subjects who received systemic non-protocol anticancer therapy (sNPACT) after randomisation and before experiencing an event, were right censored at the date of the last ATA on or prior to the date of initiation of sNPACT.

- Subjects who had not experienced an event (and were not otherwise censored) at the time of data cutoff, were right censored on the date of their last ATA prior to the cutoff date or on the date of randomisation if they had no post baseline ATAs.

- Subjects who experienced an interval corresponding to 2 or more consecutive scheduled tumour assessments (operationally defined as an interval > 178 days)

without an ATA immediately followed by an event were right censored on the date of their most recent ATA prior to the missing/inadequate assessments. However, if such an interval was immediately followed by an ATA with an overall response of SD, PR, or CR, this was deemed sufficient clinical evidence that progression did not occur during the period of missing data and the missing evaluations were ignored.

Alliance censoring rules were the following:

- Subjects who did not experience an event at the time of data cutoff were right censored on the date of their last tumour assessment; subjects with no post-baseline tumor assessment were censored at the date of randomisation. Censoring was not imposed for sNPACT or for missing or inadequate tumour assessment

A7.

According to Table 12 of the company submission, there was an imbalance between the cabozantinib and sunitinib arms in (i) the proportions of patients who had ≥2 missed “adequate tumour assessments” before a PFS event, and (ii) the proportions who had no post-baseline “adequate tumour assessments”. What are the reasons for these differences?

Response: Information on the reasons for the differences between the above imbalances is not available.
Sensitivity analysis on the proportions of patients who had ≥2 missed “adequate tumour assessments” before a PFS event has, however, been performed during the regulatory process in line with EMA guidelines on censoring rules. In this analysis the patients were counted as having events using the first date when there is no documented evidence that the criteria have been met. The results are consistent with those presented in Table 12 of our submission document. The HR adjusted for stratification factors was 0.48 and the median duration of PFS was estimated to be 8.6 months in the cabozantinib arm vs. 5.3 months in the sunitinib (see the below table).

Comparison of results of PFS Analyses when Subjects with 2 or more missed ATAs prior to event were censored (FDA censoring rules) or were not censored (EMA guideline EMA/CHMP/27994/2008/Rev.1)

	CABOMETYX (N=79)	Sunitinib (N=78)
Analysis per FDA censoring rules		
Progression-free survival (PFS) by IRC		
Median PFS in months (95% CI)	8.6 (6.8, 14.0)	5.3 (3.0, 8.2)
HR (95% CI); stratified [a,b]	0.48 (0.31, 0.74)	
Two-sided log-rank p-value: stratified [a]	p=0.0008	
Analysis per EMA Guideline (EMA/CHMP/27994/2008/Rev.1) for missed visits		
Progression-free survival (PFS) by IRC:		
Median PFS in months (95% CI)	8.6 (6.2, 14.0)	5.3 (3.0, 8.2)
HR (95% CI); stratified [a,b]	0.48 (0.32, 0.73)	
Two-sided log-rank p-value: stratified [a]	p=0.0005	

A8.

Section B.2.6 of the company submission states that, for the IRC assessment of PFS, 13 patients did not have complete data for radiographic images or tumour assessments. Figure 8, showing the percentage change in tumour size, states that 24 patients did not have adequate post-baseline imaging assessments. Please explain the difference in the number of missing observations between these two outcomes.

Response: The difference in patient numbers is due to the timing of the assessments for ORR responses being the entire period prior to progression, while for PFS only the response at time of progression is considered. The difference in the patient numbers seen thus reflects the fact that ORR and PFS are mostly evaluated at different numbers of points. To clarify further, for ORR a patient may have missing data for certain time points, leading to not having adequate post-baseline imaging assessments for evaluation of response, but still have adequate data for the single time point of progression.

To confirm numbers:

- For the trial, radiographic images were collected for 156 of 157 patients
- For PFS, a complete set of baseline and post-baseline radiographic images was available for 143 patients. The remaining 13 patients were not assessed due to the reasons given below:
    - 11 patients baseline images available but no post-baseline tumour assessments performed (1 cabozantinib, 10 sunitinib),
    - 1 patient (cabozantinib) incomplete set of baseline image collected
    - 1 patient (sunitinib) corrupt baseline image

- For ORR, and as per Table 27 of the CSR, 24 patients were not assessed (6 cabozantinib and 18 sunitinib) due to the reasons below:
    - 10 patients unable to evaluate (4 cabozantinib, 6 sunitinib)
    - 14 patients with missing data (2 cabozantinib, 12 sunitinib). This included*:
        - 2 patients (sunitinib) no baseline or post-baseline images
        - 14 patients no post-baseline images (2 cabozantinib, 12 sunitinib)

*Patients could be counted more than once

A9.

Subgroup analyses.

a) Was any adjustment made for multiple testing among the 16 subgroup analyses reported in Figure 53 in the Appendix?

Response: We can confirm that no adjustment was made.

b) Please supply the subgroup analyses results for OS.

Response: The Kaplan-Meier analysis of OS by Subgroup (ITT population is provided below.

	Cabozantinib (N=79)			Sunitinib (N=78)					
	n	Events (%)	Median (95% CI)	n	Events (%)	Median (95% CI)	Logrank p-value [1]	HR (95% CI) [2]	p-val[2] for trt*subgroup interaction
Overall									
Unstratified	79	38 (48%)	30.3 (14.6, NE)	78	45 (58%)	21.0 (16.3, 27.0)	0.1921	0.75 (0.49, 1.16)	NE
Stratified							0.1700	0.74 (0.47, 1.14)	
Age Group (Years)									0.4579
<65	45	20 (44%)	NE (14.4, NE)	42	22 (52%)	26.0 (20.7, 28.1)	0.7147	0.89 (0.49, 1.64)	
>=65	34	18 (53%)	30.3 (12.7, NE)	36	23 (64%)	16.8 (8.0, 21.2)	0.1168	0.61 (0.32, 1.14)	
Sex									0.1083
Male	66	32 (48%)	30.3 (14.5, NE)	57	29 (51%)	23.5 (17.2, NE)	0.8775	0.96 (0.58, 1.59)	
Female	13	6 (46%)	35.0 (6.8, 35.0)	21	16 (76%)	10.3 (3.4, 22.4)	0.0317*	0.35 (0.13, 0.95)	
Race Group									0.2276
White	70	32 (46%)	35.0 (14.5, NE)	75	44 (59%)	21.0 (16.3, 26.0)	0.1038	0.69 (0.43, 1.08)	
Other	9	6 (67%)	22.7 (4.0, 30.3)	3	1 (33%)	NE (1.0, NE)	0.3549	2.67 (0.31, 23.07)	
Baseline ECOG									0.1175
0	36	12 (33%)	35.0 (35.0, NE)	36	18 (50%)	26.0 (16.8, NE)	0.1055	0.54 (0.26, 1.15)	
1	33	19 (58%)	19.0 (12.7, 30.3)	32	23 (72%)	16.3 (8.0, 21.8)	0.1798	0.66 (0.36, 1.21)	
2	10	7 (70%)	11.1 (1.1, NE)	10	4 (40%)	NE (2.2, NE)	0.1815	2.28 (0.66, 7.91)	
Baseline Bone Metastasis per IxRS									0.5705
Yes	29	16 (55%)	16.7 (10.7, NE)	28	16 (57%)	20.5 (7.5, NE)	0.7476	0.89 (0.45, 1.79)	
No	50	22 (44%)	30.3 (17.5, NE)	50	29 (58%)	23.5 (16.3, 27.4)	0.1559	0.67 (0.38, 1.17)	
Baseline Heng Risk Factors per IxRS									0.4904

	Cabozantinib (N=79)			Sunitinib (N=78)					
	n	Events (%)	Median (95% CI)	n	Events (%)	Median (95% CI)	Logrank p-value [1]	HR (95% CI) [2]	p-val[2] for trt*subgroup interaction
Intermediate	64	30 (47%)	30.3 (16.4, NE)	63	34 (54%)	23.5 (18.9, 28.1)	0.3784	0.80 (0.49, 1.31)	
Poor	15	8 (53%)	18.4 (6.1, NE)	15	11 (73%)	6.4 (2.2, 22.4)	0.1566	0.51 (0.20, 1.32)	
MET Status									0.0002*
Positive	32	12 (38%)	35.0 (26.4, NE)	30	22 (73%)	16.3 (7.5, 21.0)	0.0002*	0.27 (0.13, 0.56)	
Negative	39	20 (51%)	18.8 (13.5, NE)	30	14 (47%)	26.0 (18.9, NE)	0.4333	1.31 (0.66, 2.61)	
Missing	8	6 (75%)	7.6 (1.3, 10.5)	18	9 (50%)	27.0 (8.3, NE)	0.0219*	3.51 (1.13, 10.91)	

A10.

In relation to OS and PFS, it is stated that the plausibility of different extrapolations was assessed by visual inspection by oncologists currently practising within the NHS in England. Please can you elaborate on this process, including the number of oncologists consulted, their clinical speciality, etc.

- **Response:** Visual inspection was completed by three oncologists practising within the NHS in England and included a:
    - Professor and Consultant Medical Oncologist, specialising in Uro-Oncology
    - Consultant Medical Oncologist, specialising in Uro-Oncology as well as Gynaecology Oncology and Rare Tumours
    - Consultant Clinical Oncologist, specialising in breast and urological cancers

Face-to-face meetings were held with each clinician to take them through the economic model and the graphs of the extrapolated data. For these graphs each clinician was asked to select which they considered were the most plausible based on their clinical experience of the disease area and treatment.

A11.

On page 56 of the company submission it is stated that there was "use of two independent reviewers in the SLR". Please can you clarify which of these processes this applied to: screening titles and abstracts, screening full texts, and/or data extraction and critical appraisal of included studies?

Response: During all four processes, at least two independent reviewers were involved: Double independent record selection was performed during both screening of titles/abstracts, as well as full texts. Discrepancies were resolved after discussion between reviewers, or through reconciliation by a third reviewer. Data were extracted by one reviewer. A second reviewer independently verified the extracted data (100%). In case of disagreements, a third reviewer was involved to reach consensus. Critical appraisal of included studies was performed by two independent reviewers. In case of disagreements, a third reviewer was involved to reach consensus.

A12.

On page 56 of the company submission it is stated "Additionally, the SLR may have been impacted by the decision to exclude certain studies due to data availability". We assume that this is referring to the 9 full text articles listed in Figure 3 excluded as "article not obtained". Further, Table 8 in the appendix lists 9 references which appear to have been excluded as not obtainable. We assume that these are the same 9 listed in Figure 3. Please can you confirm this, and whether there are any other studies that were excluded due to data availability, and also what the impact may have been of any exclusions?

Response: We confirm that the 9 full text articles listed in figure 3 excluded as "Article not obtained" are the same as the articles listed in table 8 in appendix D1. They were identified through the systematic review in the pazopanib manufacturer submission document. Eight of the nine references are conference abstracts. Despite

several attempts at different times, it was not possible to download them from the American Society of Clinical Oncology website. We then examined their relevance on the basis of the abstract text provided by the database search and found that these references would have been either excluded or superseded by a more recent full text publication (as outlined in table 8 column “Final conclusion/comment”). Another citation (Negrier 2006) was mentioned in table 10 of the pazopanib manufacturer submission document but a corresponding full reference could not be identified in the bibliography of this report, whereas the study it was referring to was included in our SLR. We therefore concluded that we did not miss any relevant study or publication and the exclusion of the nine articles not obtained did not impact the results of our SLR.

A13.

Please can you confirm whether you are aware of any planned/ongoing trials or studies of cabozantinib for this indication?

Response: We are not aware of any planned or ongoing trials of cabozantinib (as a single therapy agent) for this indication.

A14.

What is included in the ‘other’ reason for discontinuation from study treatment in Figure 52 in appendix D1.2?

Response: The ‘other’ reasons for discontinuation from study treatment are as follows:

N=1 cabozantinib: Subject was hospitalised for burns and study treatment discontinued to allow the burns to heal.

N=1 sunitinib: Treatment discontinued due to the subject being referred to a hospice.

A15.

The critical appraisal of study quality (Table 15, appendix D1.3) states that there were no unexpected imbalances in dropouts in the CABOSUN trial. However, Figure 52 in appendix D1.2 shows ‘withdrawal by subject during treatment’ as more than double in the sunitinib arm (n=7) compared with the cabozantinib arm (n=3). While these numbers are low, please explain why you did not consider this an imbalance.

Response: The most frequent primary reasons for study treatment discontinuation were disease progression (56% in the cabozantinib arm and 53% in the sunitinib arm) and AEs (20% and 21%, respectively). In general, numbers of dropouts were considered balanced and the particular number of ‘withdrawal by subject during treatment’ is considered as low.

A16.

Table 26 lists solicited adverse events, and unsolicited related AEs reported by ≥10% of patients in either treatment arm. The text on page 59 lists grade 3 or 4 AEs occurring in ≥5% of patients. Some of the AEs in the text as occurring in ≥5% patients (e.g. stomatitis in the cabozantinib arm) are listed in the table as reported by

≥10% of patients in either treatment arm. Please can you explain the criteria for inclusion of adverse events on page 59 and Table 26.

Response: The ≥5% refers to grade 3 or 4 AEs occurring in ≥5% of patients, while ≥10% refers to any grade AE occurring in ≥10% of patients.

Subsequent anticancer therapies

A17.

On page 44 of the company submission, it is stated that subsequent non-radiation anticancer therapy was used by 57% and 58% of patients in the cabozantinib and sunitinib arms, respectively. Table 16 states that first systemic subsequent anticancer therapy was used in 37 (46.8%) patients in the cabozantinib arm. We presume that the figure of 57% stated on page 44 includes first and subsequent treatment lines, and/or the figures in Table 50 in Appendix L are from a different data-cut – please can you confirm?

Response: Please note that both sets of figures are all from the data cut off of 15 September 2016.

To confirm, the figure of 57% refers to the percentage of patients in the cabozantinib arm who received on-study non-protocol anticancer treatments. The figure of 46.8% in Table 16 refers to the percentage of cabozantinib patients receiving the first subsequent anticancer treatment. Only first non-protocol treatments and concomitant medications were captured in the case report form.

A18.

In Table 56 of the company submission, the proportion of patients who received nivolumab as a subsequent treatment was 13% and 15% in cabozantinib and sunitinib patients, respectively. In Appendix L (Table 50) the proportion of patients receiving nivolumab as subsequent treatment is 5% and 8% in the cabozantinib and sunitinib arms, respectively. We assume that in Table 56 of the company submission you have summed the percentage of patients receiving nivolumab with the percentage of patients who received PD-1 inhibitor from Appendix Table 50. Please can you confirm if our assumption is correct?

Response: Yes, we confirm that the percentages of nivolumab as a subsequent treatment in Table 56 are the combined percentages of patients receiving anit-PD-1/PD-L1 targeting agents from Appendix Table 50.

A19.

Please can you report the number (%) of patients who received each subsequent line of systemic anticancer therapy (e.g. second, third, fourth etc)?

Response: This information was not collected in the CABOSUN study. Only first non-protocol treatments and concomitant medications were captured in the case report forms.

A20.

In the clinical study report, it is stated that “The SAP, available in Appendix 16.1.9.1, describes the statistical analyses that were performed by Exelixis”. This appendix is not available to the ERG. Please can you supply the SAP (statistical analysis plan)?

Response: A copy of Appendix 16.1.9.1 is provided (on an AIC basis) alongside this response.

Indirect treatment comparison (ITC)

A21.

For the Ouwens et al model and the fractional polynomials model please provide more information about the Bayesian methods used including the prior probability distributions (e.g. vague, informative, non-informative, the rationale for their choice, and any sensitivity analyses on these), the likelihood distribution, the number of iterations used for burn in and inferences, and the methods for assessing convergence.

Response: All models in the network meta-analysis were indeed Bayesian and captured treatment effects using either fixed-effects or random-effects. For each case the following models were prepared: Weibull, exponential, log-normal, log-logistic, 5 1st order fractional polynomial models (P = -1, -0.5, 0, 0.5, 1) and 5 2nd order fractional polynomial models ([P1, P2] = [-0.5, 0], [-1,0], [-1,-1], [-1,0.5], [-1,1]).

Ouwens et al. model details

In the set of these parametric models, the underlying hazard rate h_{jkt} in study j for treatment k is modelled at each time t (Ouwens et al. 2010) as a function of two parameters (ν_{jk}, θ_{jk}) (with the exception of the exponential distribution where only one is necessary) and t. For example, for the Weibull distribution it is:

$$\log(h_{jkt}) = \nu_{jk} + \theta_{jk}\log(t)$$

In all types of distributions used to express hazard in the parametric models (Weibull, exponential, log-normal and log-logistic), (ν_{jk}, θ_{jk}) are expressed as:

$$\begin{pmatrix}\nu_{jk}\\ \theta_{jk}\end{pmatrix} = \begin{cases}\begin{pmatrix}\mu_{1jb}\\ \mu_{2jb}\end{pmatrix}, & \text{if } k = b = \{\text{sunitinib}\}\ (\text{"baseline" treatment})\\ \begin{pmatrix}\mu_{1jb}\\ \mu_{2jb}\end{pmatrix} + \begin{pmatrix}\sigma_{1jbk}\\ \sigma_{2jbk}\end{pmatrix}, & \text{if } k \text{ different from } b\end{cases}$$

In the case of the random-effects model:

$$\begin{pmatrix}\delta_{1jbk}\\ \delta_{2jbk}\end{pmatrix} \sim N\left(\begin{pmatrix}d_{1Ak}\\ d_{2Ak}\end{pmatrix} - \begin{pmatrix}d_{1Ab}\\ d_{2Ab}\end{pmatrix}, \Sigma\right),\ \Sigma = \begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix}$$

In the fixed effects model, this last expression is replaced by:

$$\begin{pmatrix}\delta_{1jbk}\\ \delta_{2jbk}\end{pmatrix} - \begin{pmatrix}d_{1Ak}\\ d_{2Ak}\end{pmatrix} - \begin{pmatrix}d_{1Ab}\\ d_{2Ab}\end{pmatrix}$$

The vectors $\begin{pmatrix}\mu_{1|b}\\ \mu_{2|b}\end{pmatrix}$ are the treatment-specific parameters ν and θ of the reference treatment in study j. In our case study, sunitinib was chosen as the reference treatment in both CABOSUN and COMPARZ, a choice that does not impact the final results (see Ouwens et al. 2010). The vector $\begin{pmatrix}\delta_{1|bk}\\ \delta_{2|bk}\end{pmatrix}$ is the study-specific difference in scale ν and shape θ of the log-hazard curve for treatment k relative to the reference treatment in study j. In our case study, treatment k corresponded to cabozantinib in CABOSUN and pazopanib in COMPARZ.

The estimation of model parameters– reference and effect vectors – was performed under a Bayesian framework. Prior distributions used for the parameters of all models were chosen with high variability to ensure they would be non-informative, i.e., not inform or influence the result of the parameter estimation:

$$\begin{pmatrix}\mu_{1|b}\\ \mu_{2|b}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\ 0\end{pmatrix}, T_\mu\right), T_\mu = \begin{pmatrix}10^4 & 0\\ 0 & 10^4\end{pmatrix}$$

$$\begin{pmatrix}d_{1Ak}\\ d_{2Ak}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\ 0\end{pmatrix}, T_d\right), T_d = \begin{pmatrix}10^4 & 0\\ 0 & 10^4\end{pmatrix}$$

$$\Sigma \sim \text{Wishart}(\Omega, 2), \qquad \Omega = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

In the fixed effects models, the prior distribution for Σ did not need to be specified.

The default precision value of 10^4 above was chosen to be >>1, and identical to the one proposed in the original paper by Ouwens et al 2010. For models where very large initial values resulted in numerical overflow, precision was increased up to $5*10^3$. A justification of the choice of distribution for Σ can be found in Gelman (2006)[1].

The model parameters were estimated using a Markov Chain Monte Carlo (MCMC) Gibbs algorithm implemented in WinBUGS. The WinBUGS sampler, using three parallel chains, was run for 50,000 iterations as burn-in period and further 100,000 iterations for inferences for fixed effects models. Those numbers were increased to 150,000 and 200,000, respectively for the random effects models. The Gelman-Rubin statistic Rhat was calculated and convergence declared when Rhat < 1.05.

Fractional polynomial (FP) model details

As an extension to the Ouwens et al (2010) method, Jansen (2011) proposed NMA models using parametric survival functions, which include not only common survival distributions such as Weibull or Gompertz, but also more flexible fractional polynomials.

The first order fractional polynomial is written as:

$$\log(h_{jkt}) = \beta_{0jk} + \beta_{1jk} t^p, \text{ with } t^0 = \log(t) \text{ and}$$

$$\begin{pmatrix}\beta_{0|k}\\ \beta_{1|k}\end{pmatrix} = \begin{cases}\begin{pmatrix}\mu_{0|b}\\ \mu_{1|b}\end{pmatrix}, & \text{if } k = b, b = \{\text{sunitinib}\}\\ \begin{pmatrix}\mu_{0|b}\\ \mu_{1|b}\end{pmatrix} + \begin{pmatrix}d_{0Ak} - d_{0Ab}\\ d_{1Ak} \quad d_{1Ab}\end{pmatrix}, & \text{if } k \text{ different from } b\end{cases}$$

[1] http://www.stat.columbia.edu/~gelman/research/published/taumain.pdf

where h_{jkt} is the underlying hazard rate in trial j for intervention k, at time t, and is now described as a function of time t to the power P.

The vector $\begin{pmatrix}\mu_{0jb}\\ \mu_{1jb}\end{pmatrix}$ pertains to treatment, whereas d_{0Ak} corresponds to the treatment effect of k relative to overall reference treatment A, and the vector $\begin{pmatrix}d_{0Ak}-d_{0Ab}\\ d_{1Ak}-d_{1Ab}\end{pmatrix}$ reflects the difference in β_{0jk} and β_1 of the log hazard curve for treatment b relative to k. Note that, if $\beta_{1jk}\neq 0$ and P = 1, a linear hazard function is obtained, which corresponds to a Gompertz survival function, and if $\beta_{1jk} \neq 0$ and P = 0, a Weibull hazard function is obtained. As such, the log-hazard function of the Weibull and Gompertz survival distributions are special cases of the fractional polynomial models.

Fixed- and random-effects models only differ, as above for the Ouwens models, by the assumption made on the distribution of $\begin{pmatrix}d_{0Ak}-d_{0Ab}\\ d_{1Ak}-d_{1Ab}\end{pmatrix}$.

For additional flexibility, this first-order fractional polynomial model can be generalised to a 2nd order fractional polynomial.

$$\log(h_{jkt}) = \begin{cases}\beta_{0jk}+\beta_{1jk}t^{p_1}+\beta_{1jk}t^{p_2}, & p_1 \neq p_2\\ \beta_{0jk}+\beta_{1jk}t^{p}+\beta_{1jk}t^{p}(\log t), & p_1 = p_2 = p\end{cases} \quad \text{with } t^0 = \log(t)$$

$$\begin{pmatrix}\beta_{0jk}\\ \beta_{1jk}\\ \beta_{2jk}\end{pmatrix} = \begin{cases}\begin{pmatrix}\mu_{0jb}\\ \mu_{1jb}\\ \mu_{2jb}\end{pmatrix}, & \text{if } k = b, b = \{\text{sunitinib}\}\\ \begin{pmatrix}\mu_{0jb}\\ \mu_{1jb}\\ \mu_{2jb}\end{pmatrix}+\begin{pmatrix}d_{0Ak}-d_{0Ab}\\ d_{1Ak}-d_{1Ab}\\ d_{2Ak}-d_{2Ab}\end{pmatrix}, & \text{if } k \text{ different from } b\end{cases}$$

Estimation of model parameters– reference and effect vectors – was performed in Bayesian framework. The prior distributions used for the parameters for the 1st order FP were chosen non-informative as follows:

$$\begin{pmatrix}\mu_{1jb}\\ \mu_{2jb}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\ 0\end{pmatrix}, T_\mu\right), T_\mu = \begin{pmatrix}10^4 & 0\\ 0 & 10^4\end{pmatrix}$$

$$\begin{pmatrix}d_{1Ak}\\ d_{2Ak}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\ 0\end{pmatrix}, T_d\right), T_d = \begin{pmatrix}10^4 & 0\\ 0 & 10^4\end{pmatrix}$$

And for the 2nd order FP:

$$\begin{pmatrix}\mu_{1jb}\\ \mu_{2jb}\\ \mu_{3jb}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}, T_\mu\right), T_\mu = \begin{pmatrix}10^4 & 0 & 0\\ 0 & 10^4 & 0\\ 0 & 0 & 10^4\end{pmatrix}$$

$$\begin{pmatrix}d_{1Ak}\\ d_{2Ak}\\ d_{3Ak}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}, T_d\right), T_d = \begin{pmatrix}10^4 & 0 & 0\\ 0 & 10^4 & 0\\ 0 & 0 & 10^4\end{pmatrix}$$

As for the Ouwens method, the default precision value of 10^4 was chosen to be >>1, and identical to the one proposed in the original paper on fractional polynomials Janssen 2011. For random effect models, $\Sigma \sim \text{Wishart}(\Omega, 2)$, $\Omega = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. A justification of this choice can be found in Gelman (2006)[2].

[2] http://www.stat.columbia.edu/~gelman/research/published/taumain.pdf

The model parameters were estimated using a Markov Chain Monte Carlo (MCMC) Gibbs algorithm implemented in WinBUGS. The WinBUGS sampler, using three parallel chains, was run for 250,000 iterations as burn-in period and further 250,000 iterations for inference in the FP model case. The Gelman-Rubin statistic Rhat was calculated and convergence declared when Rhat < 1.05.

A22.

Priority question. Please clarify the interval used for dividing up the follow-up period in the fractional polynomial model. Please supply the tabulated hazard ratios and 95% credible intervals for each fractional polynomial model for each interval time point. Please also supply hazard ratio plots for each fractional polynomial model with credible intervals to allow visual inspection.

Response: To produce a continuous curve, each curve was divided into 800 points. Since Since providing all HRs would produce a very large number of results, we have provided provided tables with monthly HRs (from Figure 11 to Table ***8***) and HR plots instead (Figure 11 to Figure 30). The 95% credible intervals were derived from 10,000 draws from the posterior distribution of hazard functions and are presented in the graphs (shaded areas) and tables. Statistical fits for all the above models are provided in Table ***9***.

OS fractional polynomial 1st order

Figure 11 Hazard ratio plot, OS; fractional polynomial 1st order (p=0)

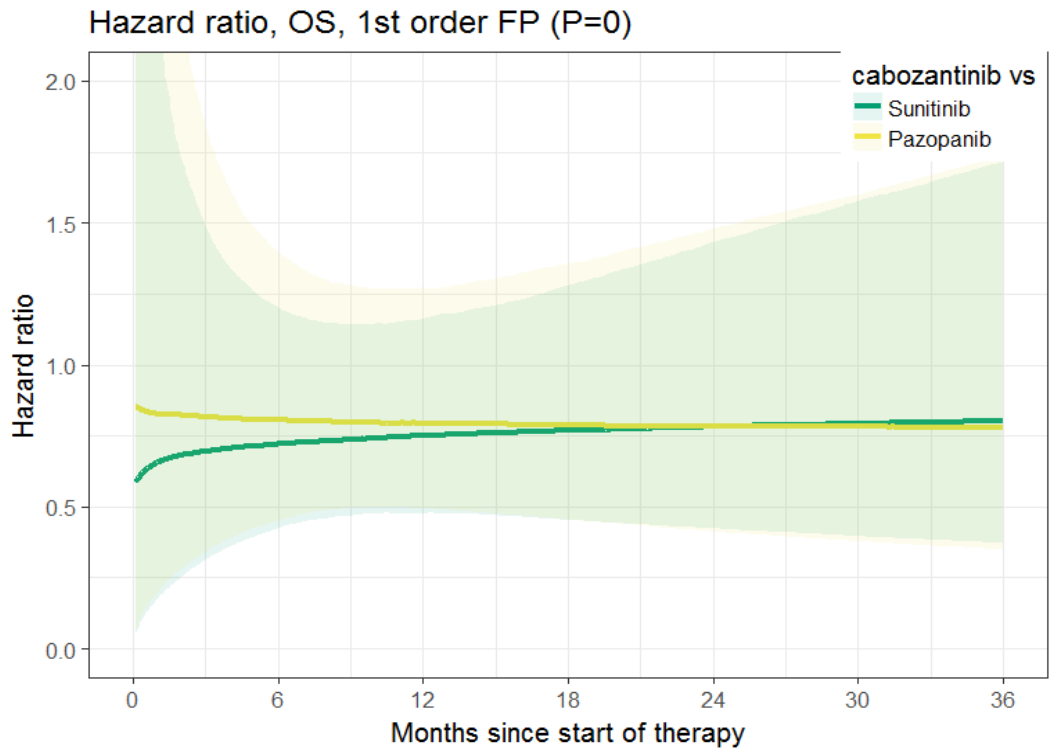


Figure 12 Hazard ratio plot, OS; fractional polynomial 1st order (p=1)

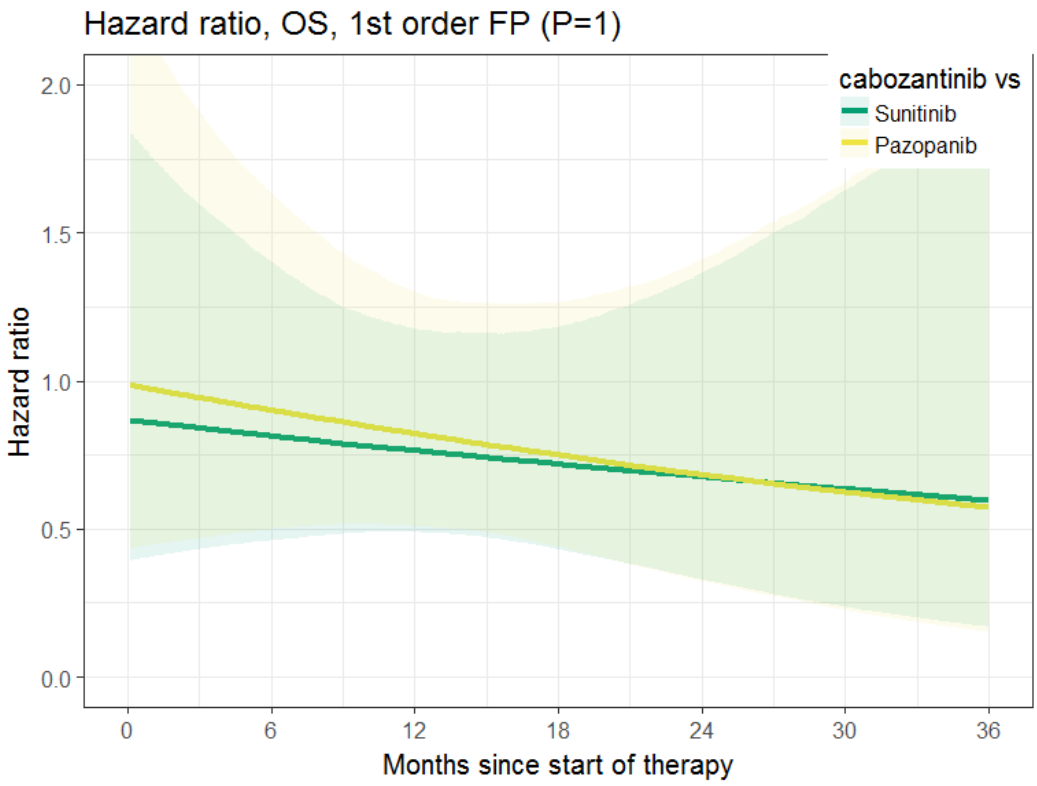


Figure 13 Hazard ratio plot, OS; fractional polynomial 1st order (p=0.5)

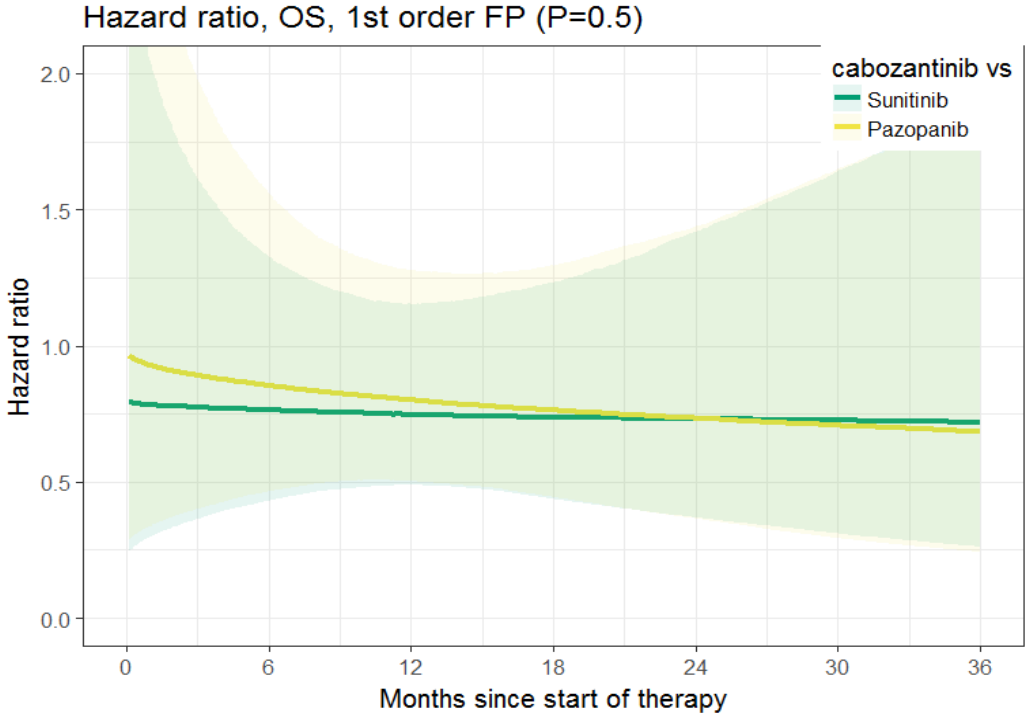


Figure 14 Hazard ratio plot, OS; fractional polynomial 1st order (p=-1)

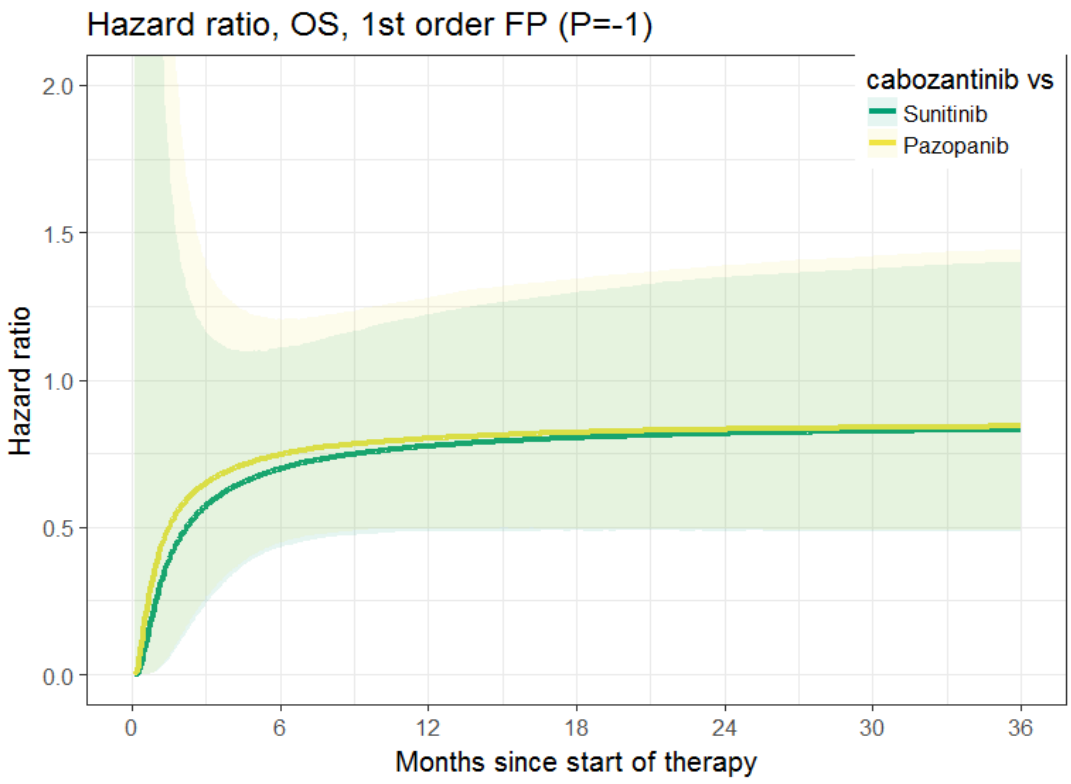


Figure 15 Hazard ratio plot, OS; fractional polynomial 1st order (p=-0.5)

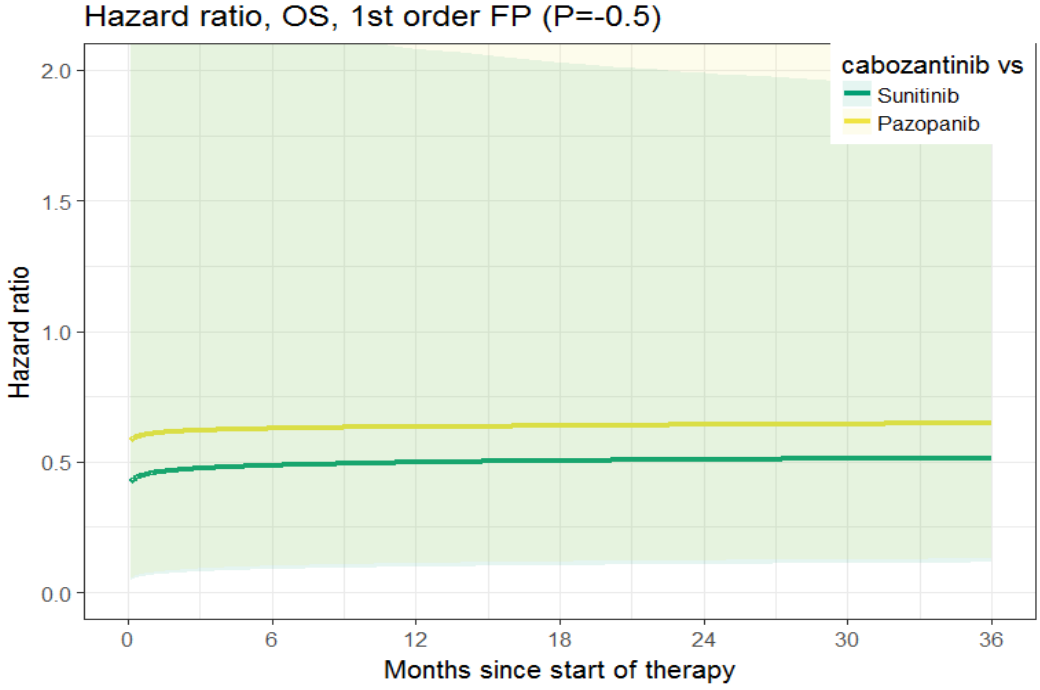


Table 1 Hazard ratios - monthly, OS fractional polynomial 1st order pazopanib

Table 2 Hazard ratios - monthly, OS fractional polynomial 1st order sunitinib

[redacted]

OS fractional polynomial 2nd order

Figure 16 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-0.5, p2=0)

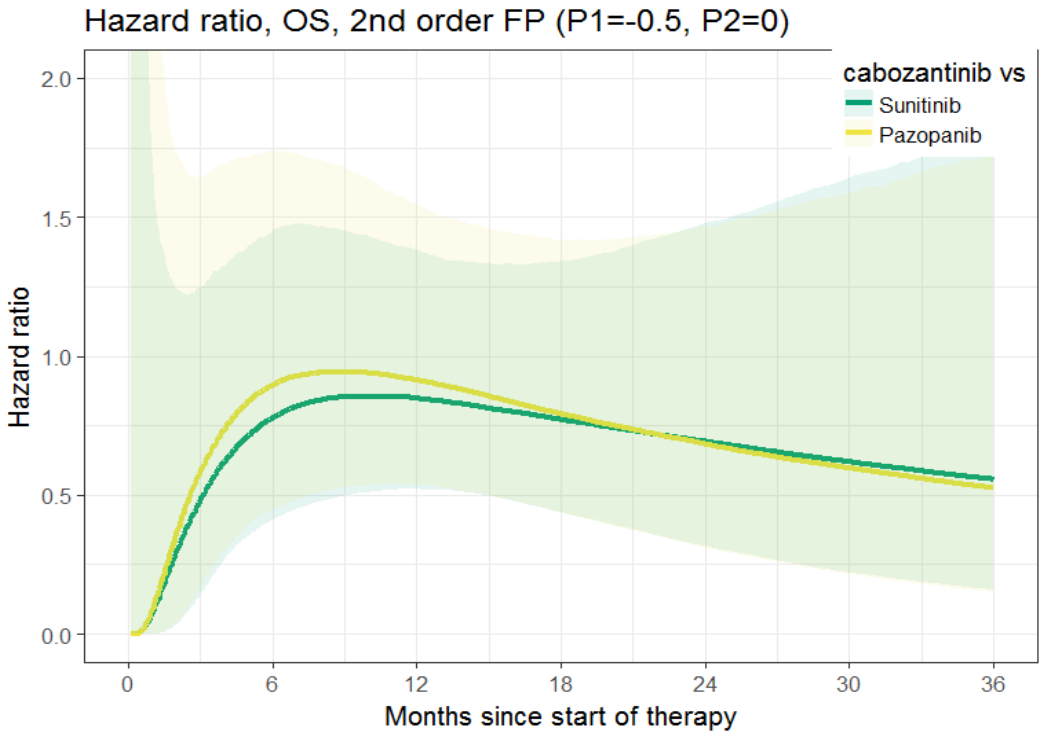


Figure 17 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, p2=0)

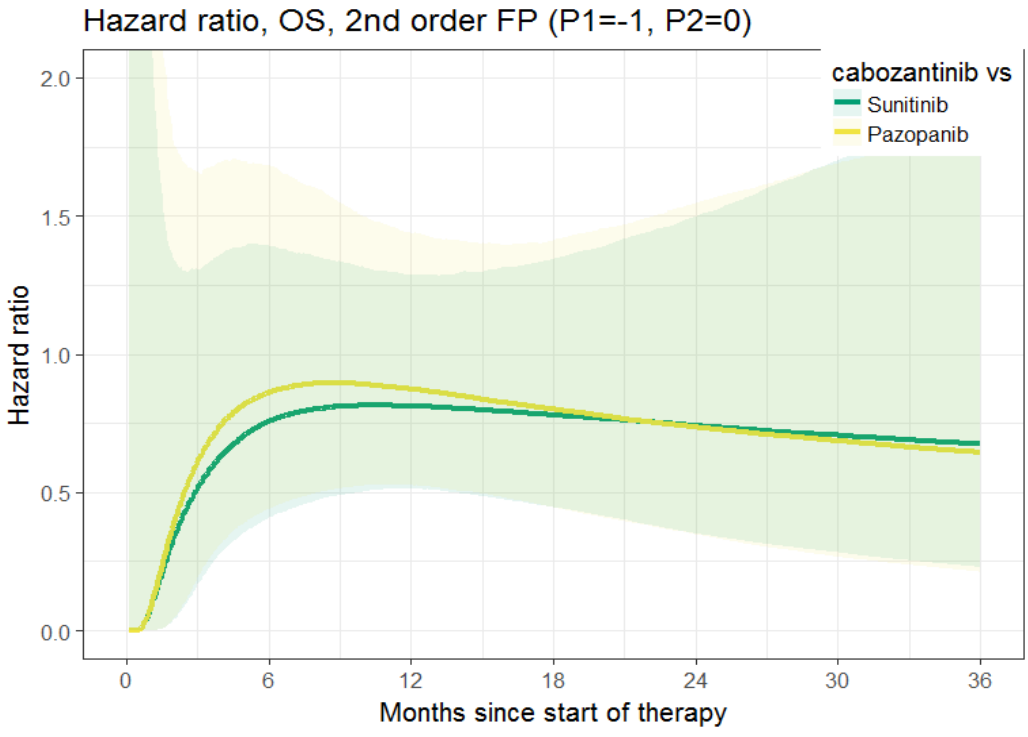


Figure 18 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, p2=-1)

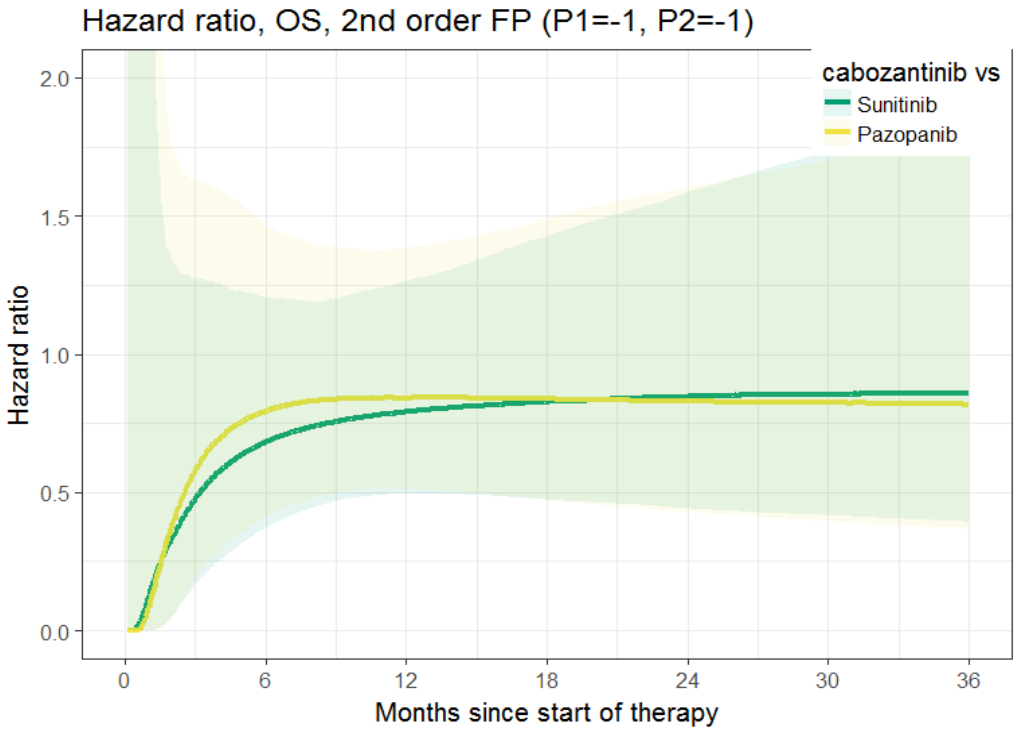


Figure 19 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, p2=0.5)

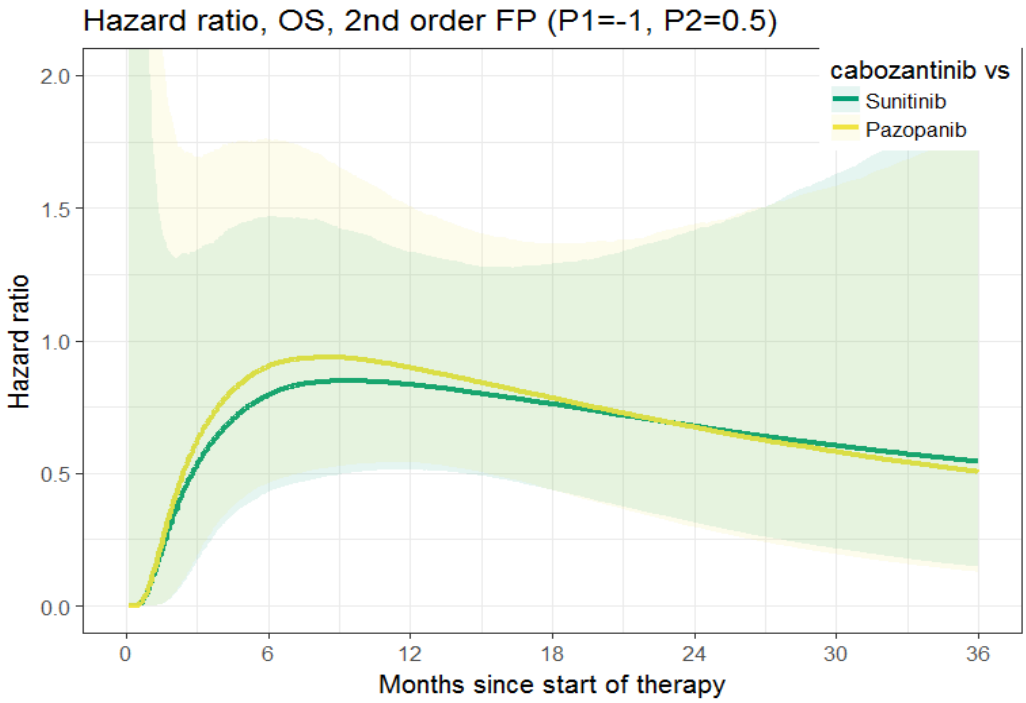


Figure 20 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, P2=1)

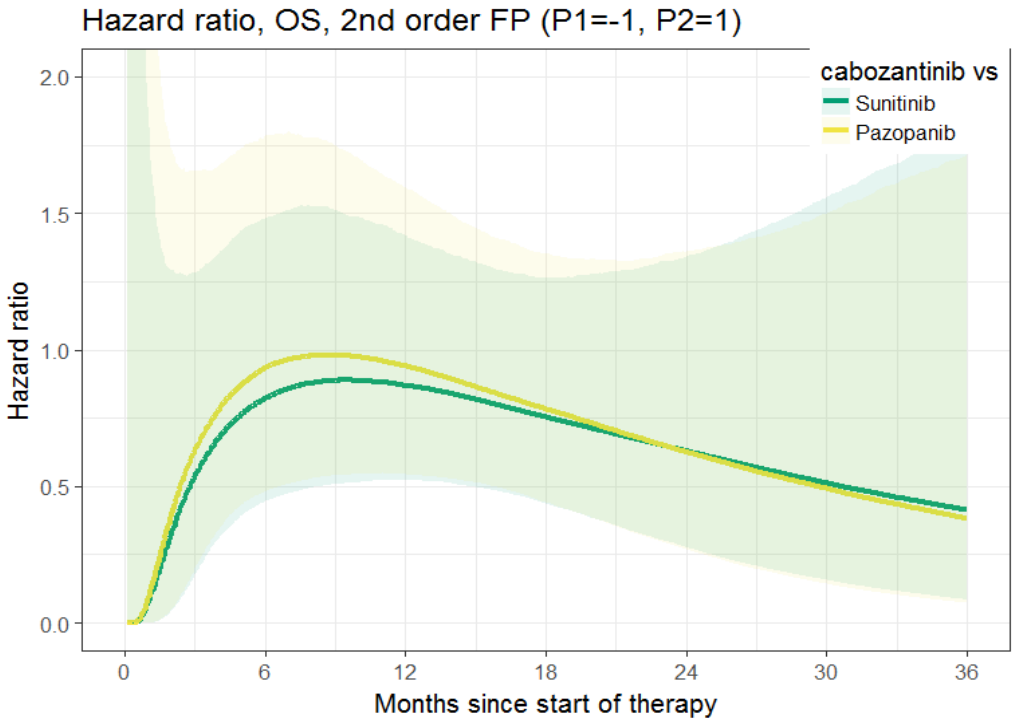


Table 3 Hazard ratios - monthly, OS fractional polynomial 2nd order pazopanib

[redacted]

Table 4 Hazard ratios - monthly, OS fractional polynomial 2nd order sunitinib

PFS fractional polynomial 1st order

Figure 21 Hazard ratio plot, PFS; fractional polynomial 1st order (p=-1)

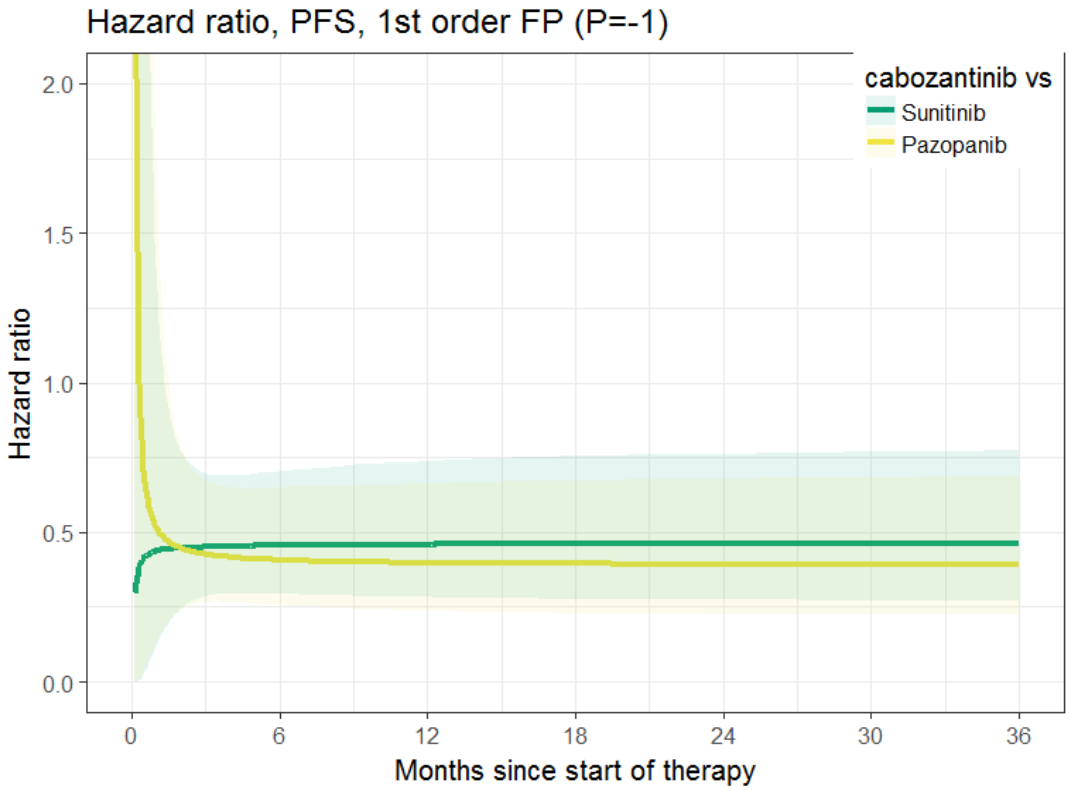


Figure 22 Hazard ratio plot, PFS; fractional polynomial 1st order (p=-0.5)

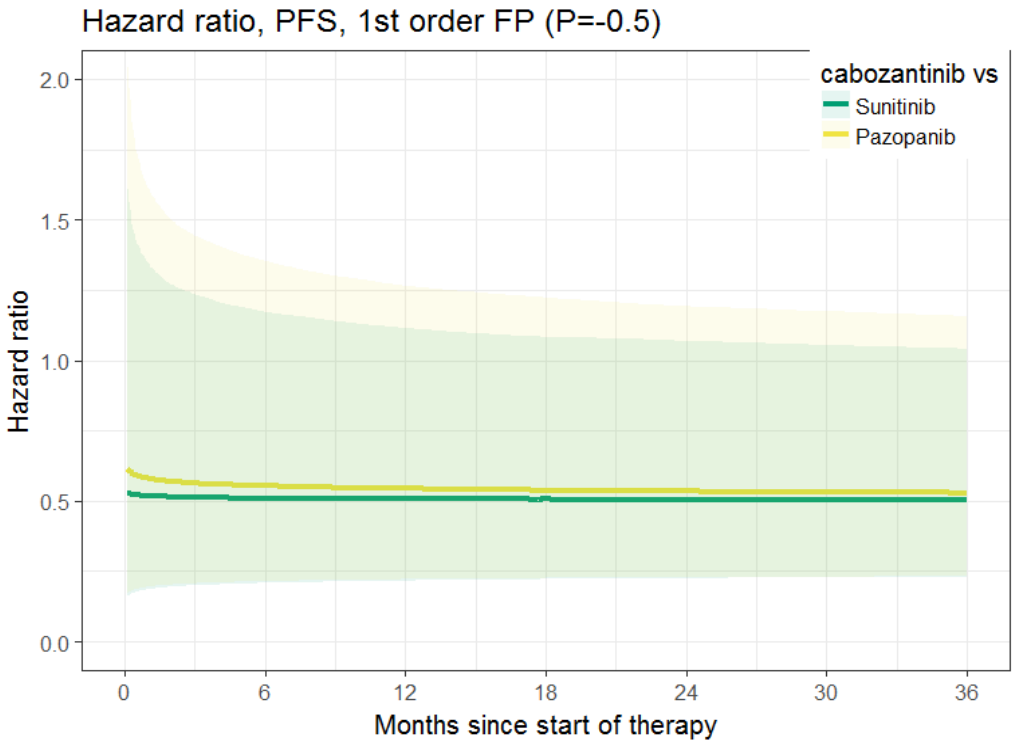


Figure 23 Hazard ratio plot, PFS; fractional polynomial 1st order (p=0)

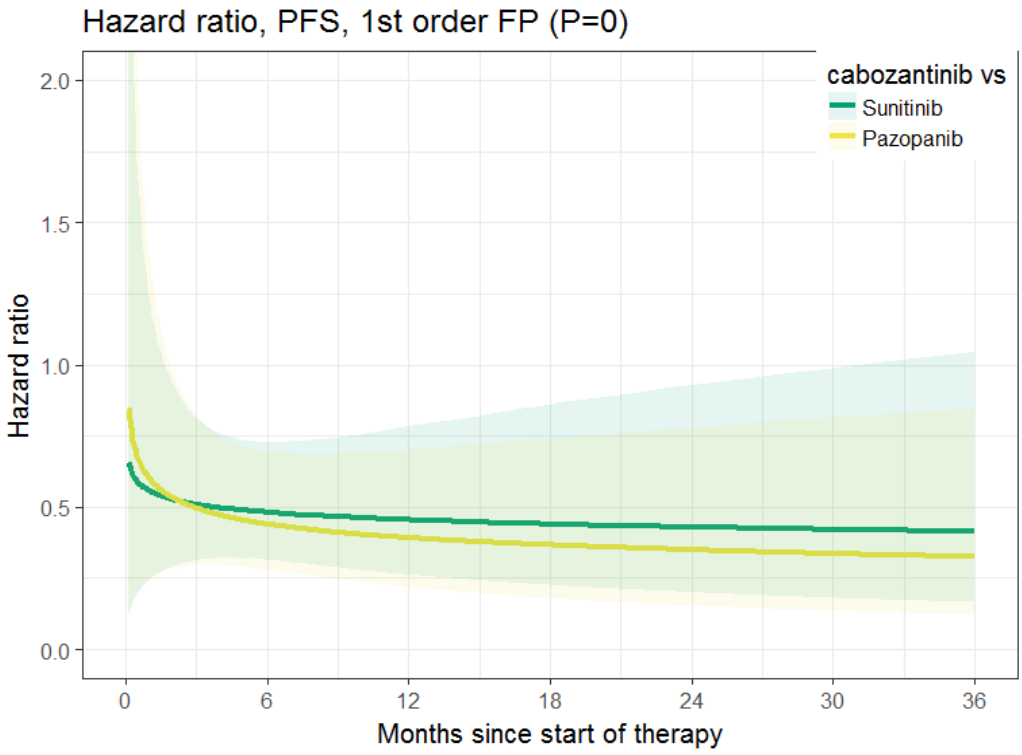


Figure 24 Hazard ratio plot, PFS; fractional polynomial 1st order (p=0.5)

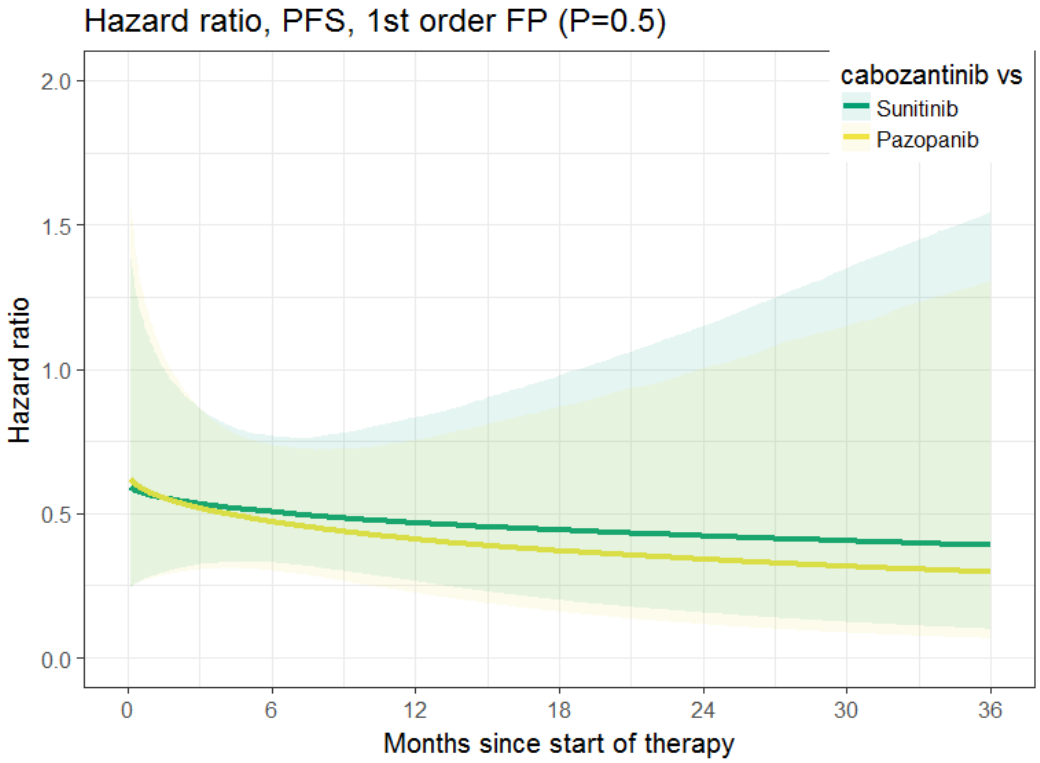


Figure 25 Hazard ratio plot, PFS; fractional polynomial 1st order (p=1)

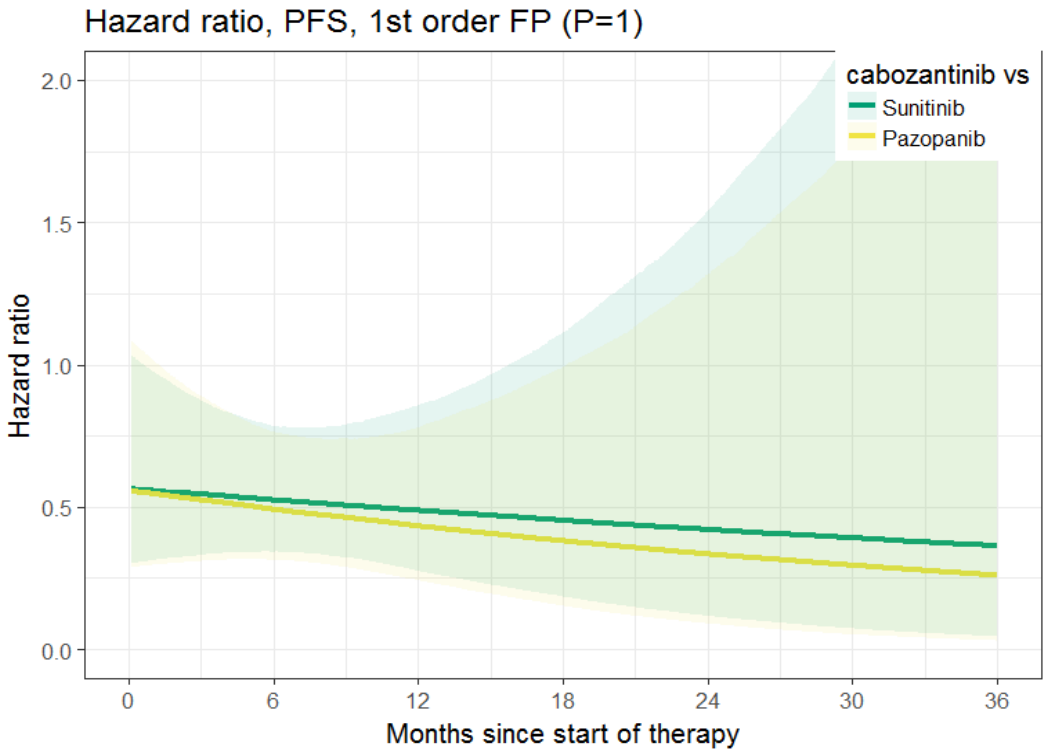
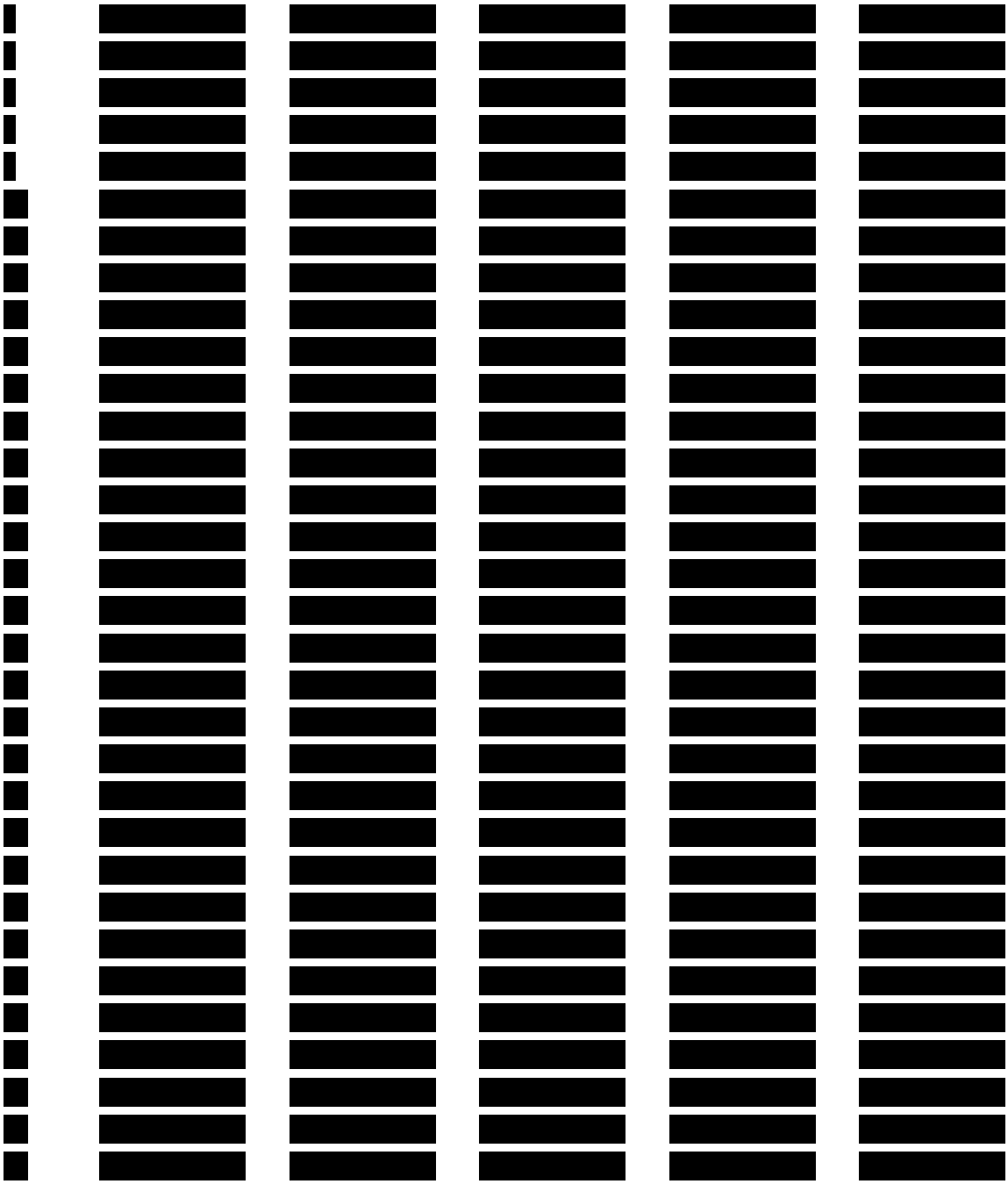


Table 5 Hazard ratios - monthly, PFS fractional polynomial 1st order pazopanib

Table 6 Hazard ratios - monthly, PFS fractional polynomial 1st order sunitinib



PFS fractional polynomial 2nd order

Figure 26 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-0.5, p2=0)

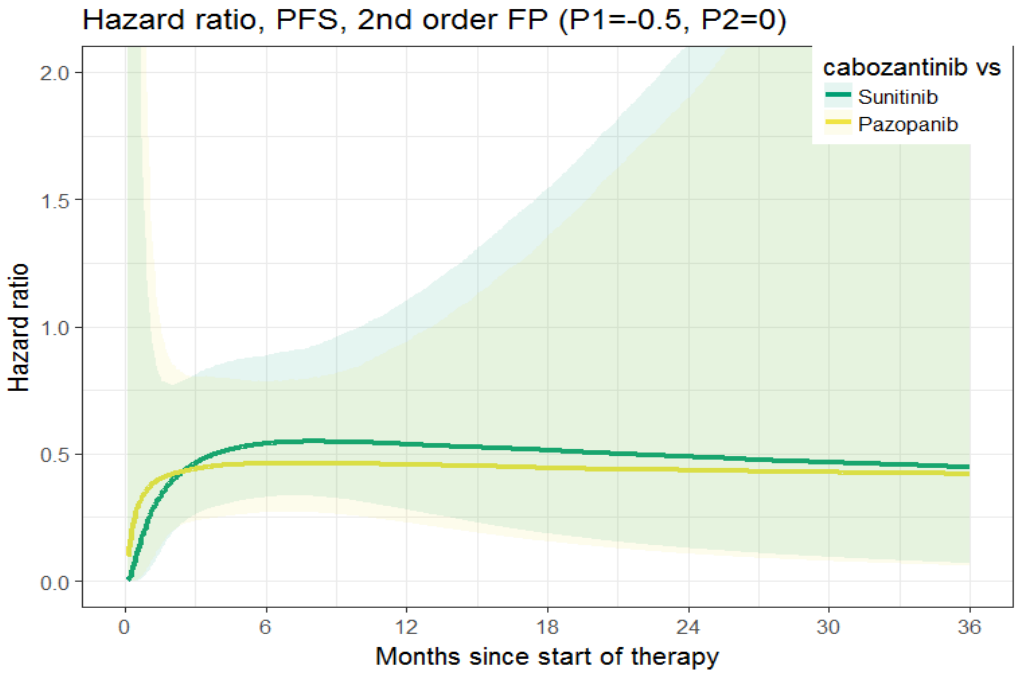


Figure 27 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=0)

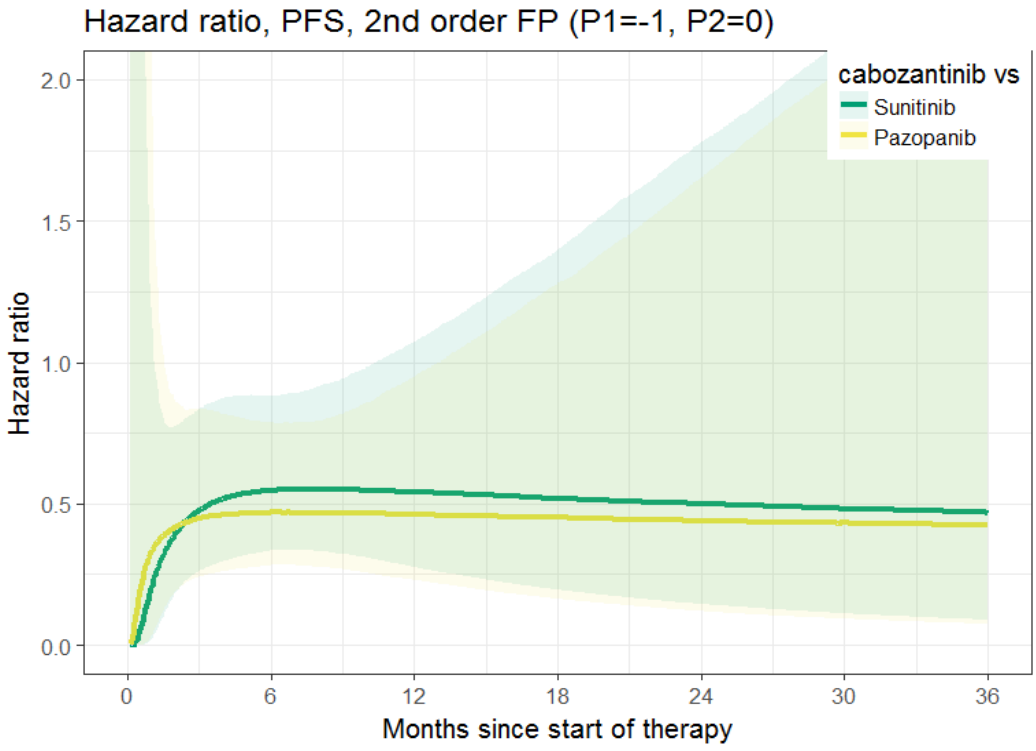


Figure 28 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=-1)

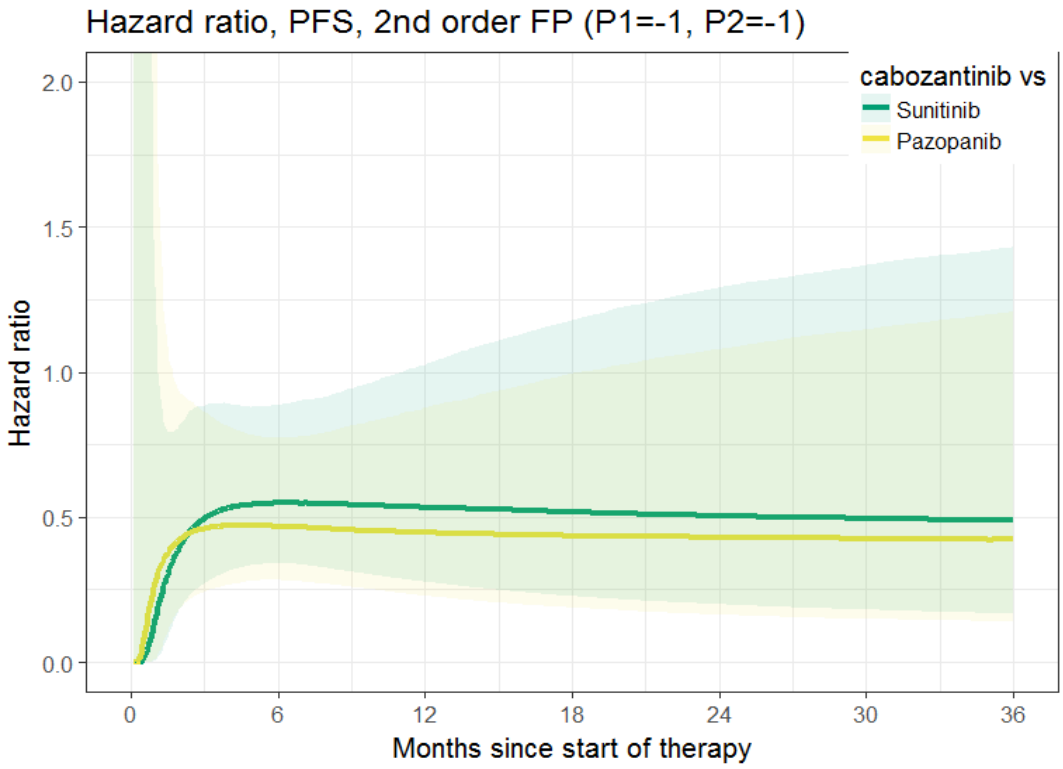


Figure 29 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=0.5)

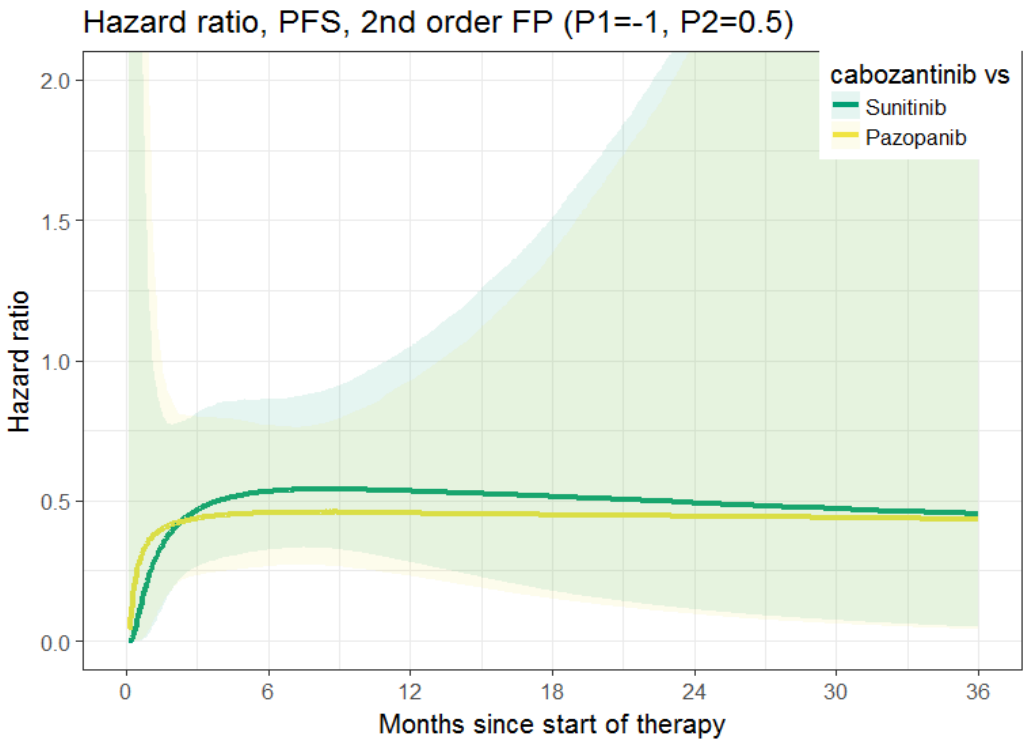


Figure 30 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=1)

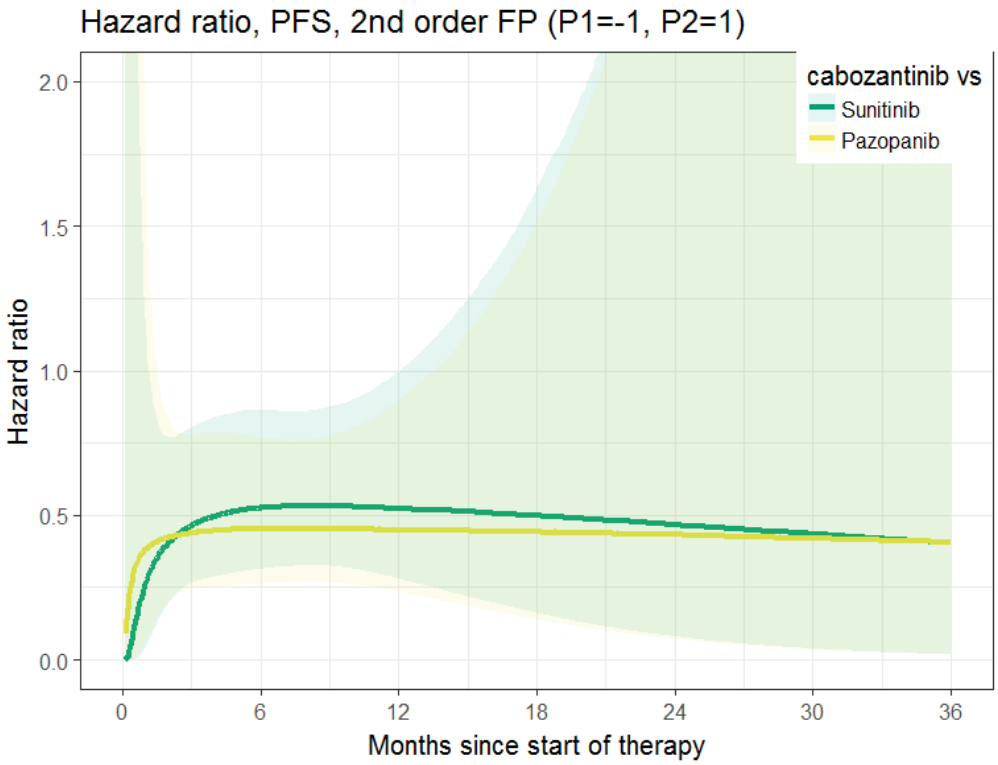


Table 7 Hazard ratios - monthly, OS fractional polynomial 2nd order pazopanib

Table 8 Hazard ratios - monthly, OS fractional polynomial 2nd order sunitinib

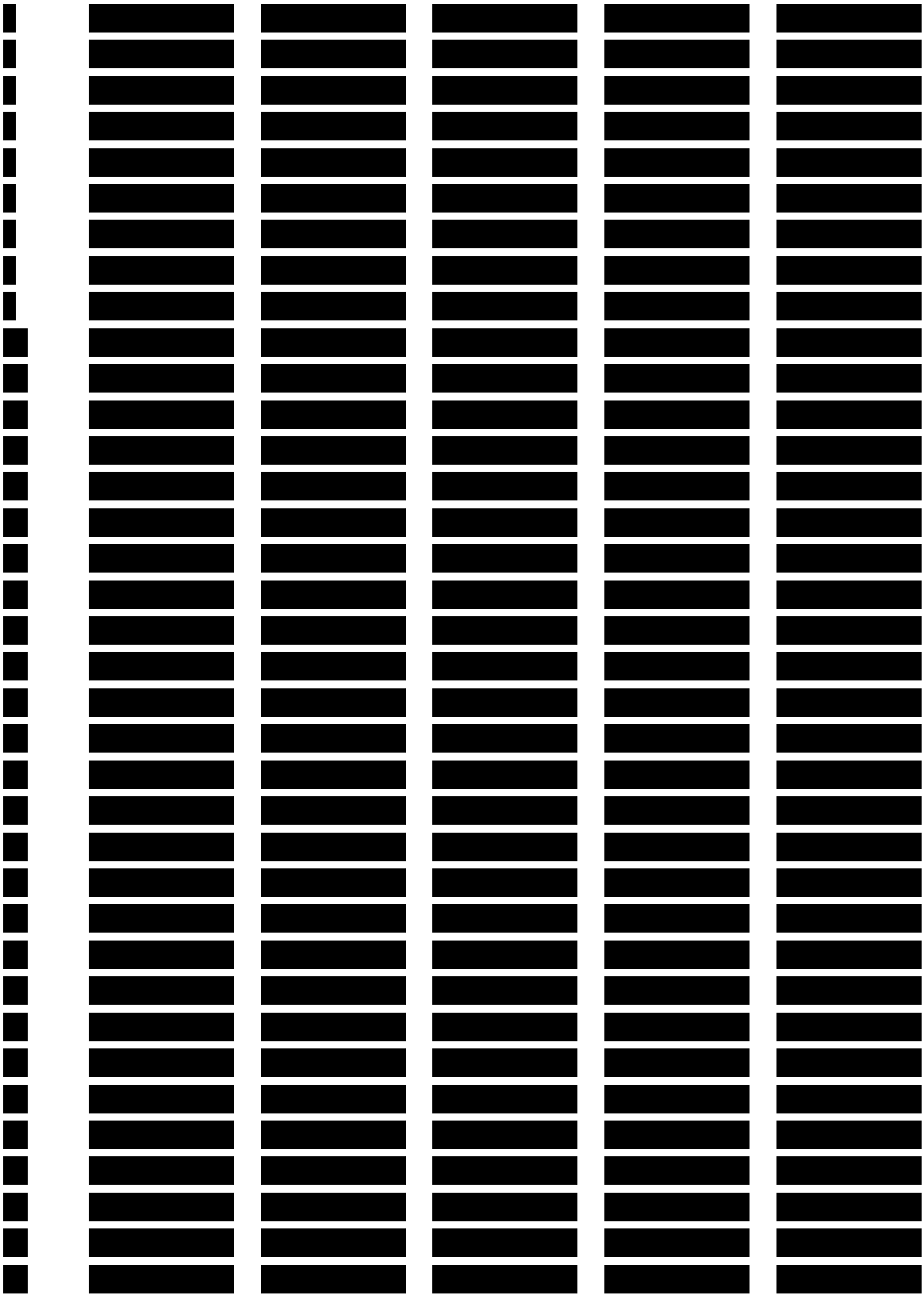


Table 9 Statistical fits – DIC (fixed effect fractional polynomial models)

FP	parameters	OS	PFS
1st	P0	1757.7	1945.5
1st	P1	1773.0	1943.6

1st	P05	1768.7	1947.4
1st	Pm1	1722.3	1910.6
1st	Pm05	1739.1	1932.4
2nd	Pm050	1717.1	1853.8
2nd	Pm10	1713.8	1840.6
2nd	Pm1m1	1712.7	1824.0
2nd	Pm105	1716.0	1850.0
2nd	Pm11	1718.8	1858.4

A23.

Priority question. Please provide fractional polynomial results (in terms of curves, and hazard ratios and 95% credible intervals for each time point) based on a random-effect model.

Response: We have run random effect models for fractional polynomials for all models submitted (Figure 31 - Figure 70). The random-effects models returned very wide 95% credible intervals for hazard ratios, OS, and PFS curves. This is due to the large variance of the treatment-effect parameters, which were all drawn from multivariate normal distributions. In the absence of multiple comparisons between treatments, which is the case for almost all treatments in the wider evidence networks and all in the small one, there is too little data to distinguish between study- and treatment-specific effects. That is especially the case when comparing cabozantinib, sunitinib and pazopanib across two studies. As a result, in random-effects models, the variance of the study-specific parameter (or random effect) is close to the variance of the treatment effect from fixed-effects models, and the variance of treatment effects ends up being very large to “accommodate” the random effect part of the model: 95% intervals for survival curves are close to (0; 1) for almost all treatments (apart from the reference treatment) and approximate (0; infinity) for hazard ratios. As the modelling objective was to infer treatment-specific effects, we only reported mean estimate for random-effects models and mean estimate with 95% intervals for fixed-effects models.

OS fractional polynomial 1st order

Figure 31 Survival curve, OS; fractional polynomial 1st order (p=0)

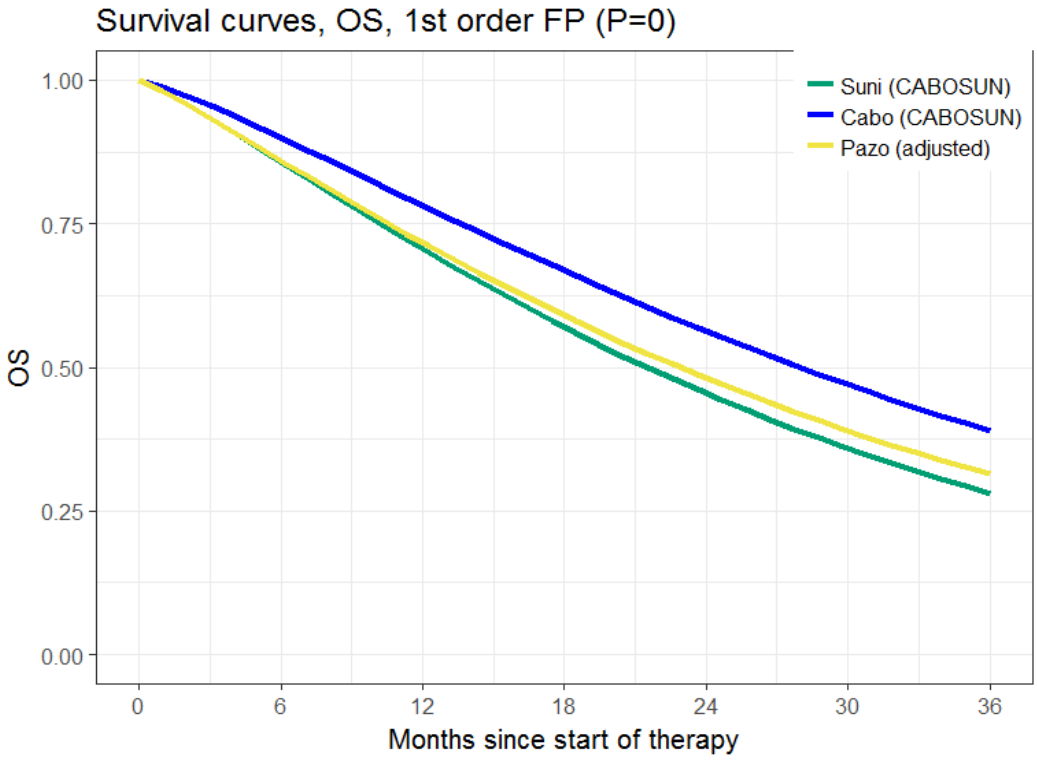


Figure 32 Hazard ratio plot, OS; fractional polynomial 1st order (p=0)

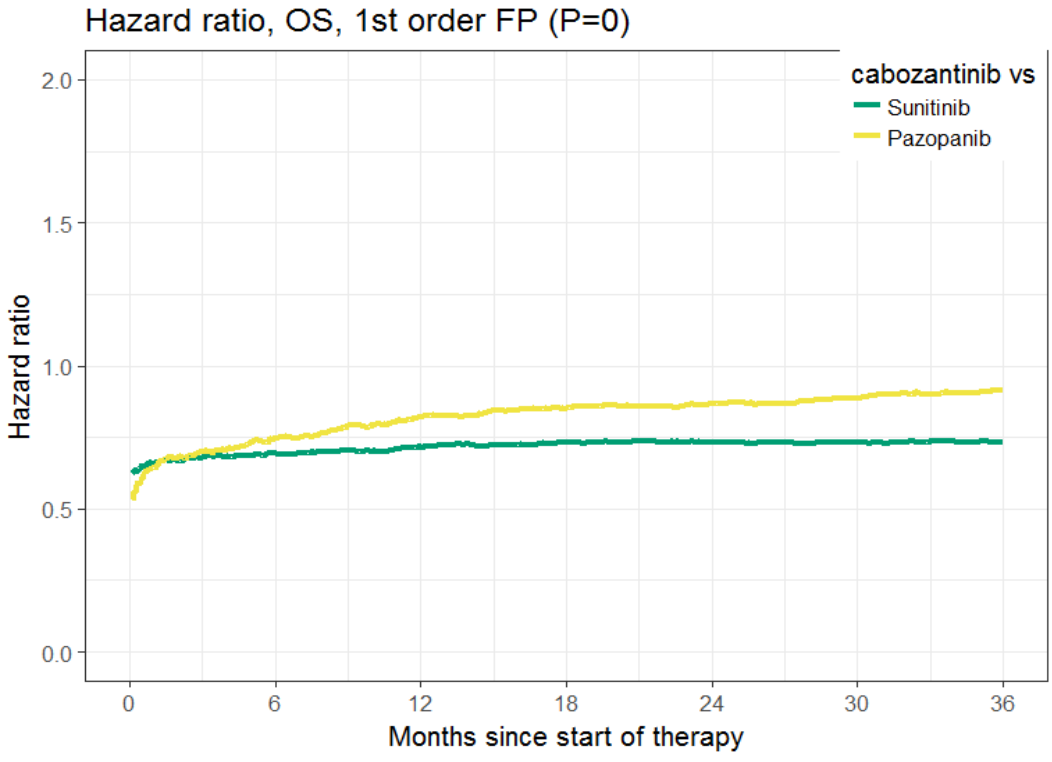


Figure 33 Survival curve, OS; fractional polynomial 1st order (p=1)

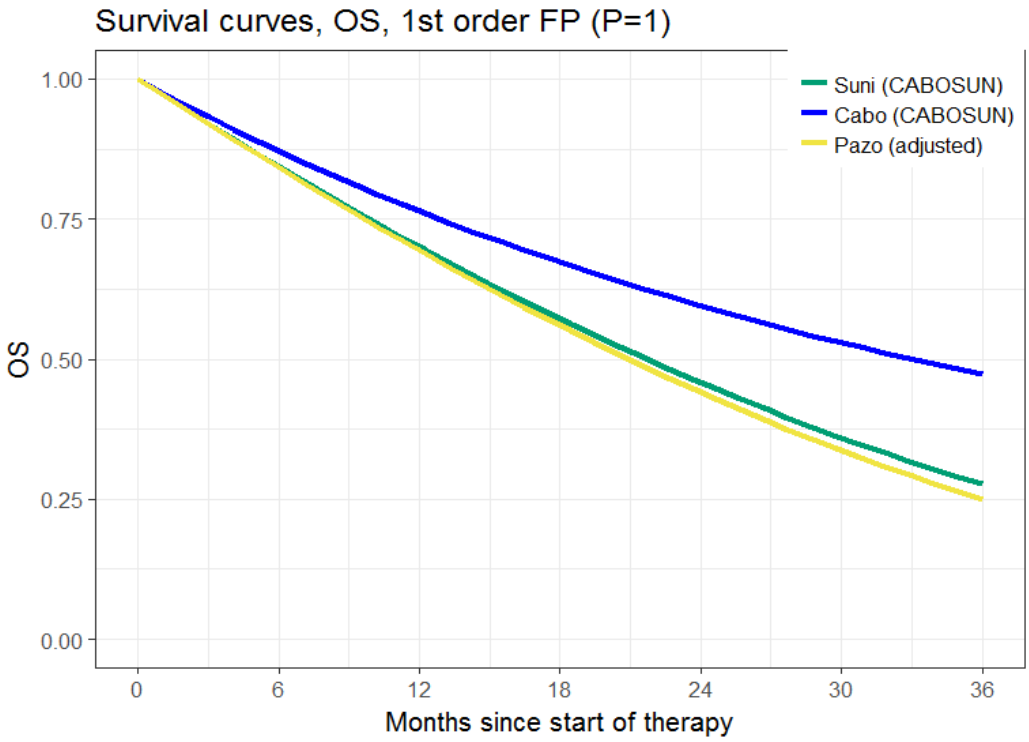


Figure 34 Hazard ratio plot, OS; fractional polynomial 1st order (p=1)

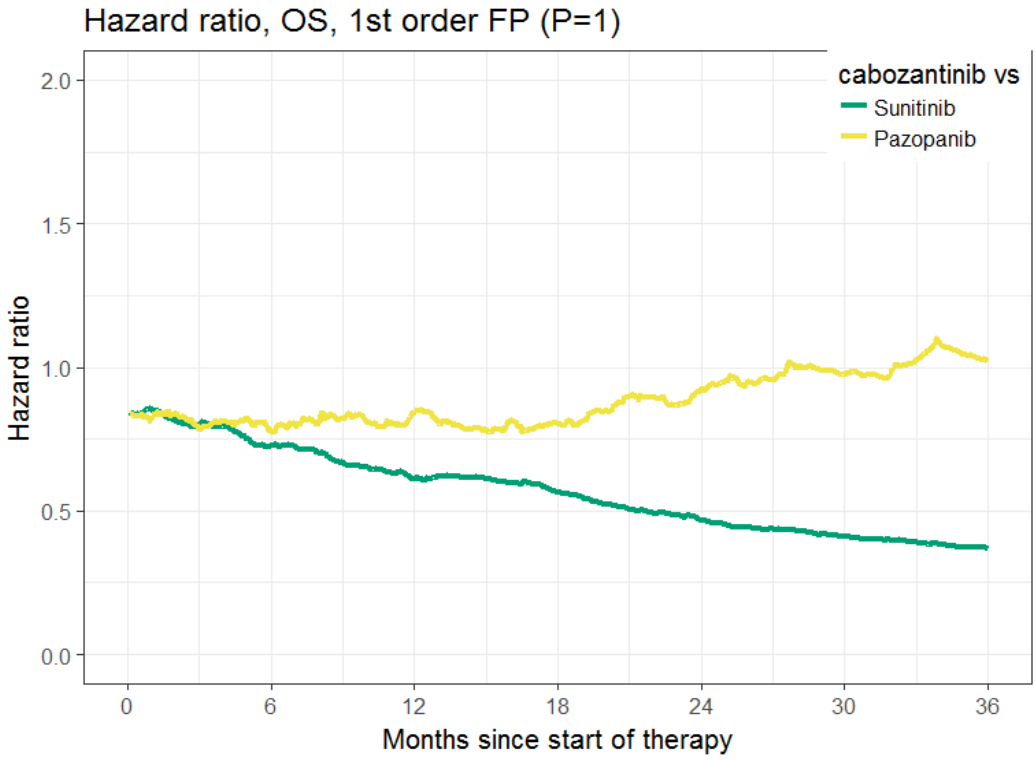


Figure 35 Survival curve, OS; fractional polynomial 1st order (p=0.5)

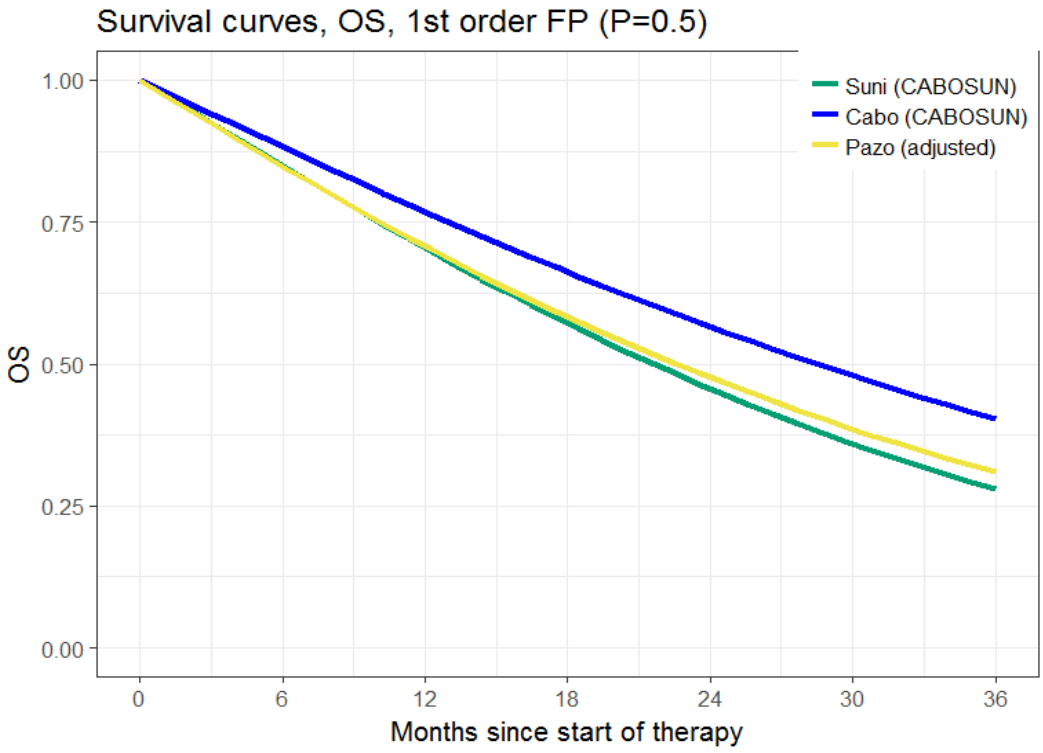


Figure 36 Hazard ratio plot, OS; fractional polynomial 1st order (p=0.5)

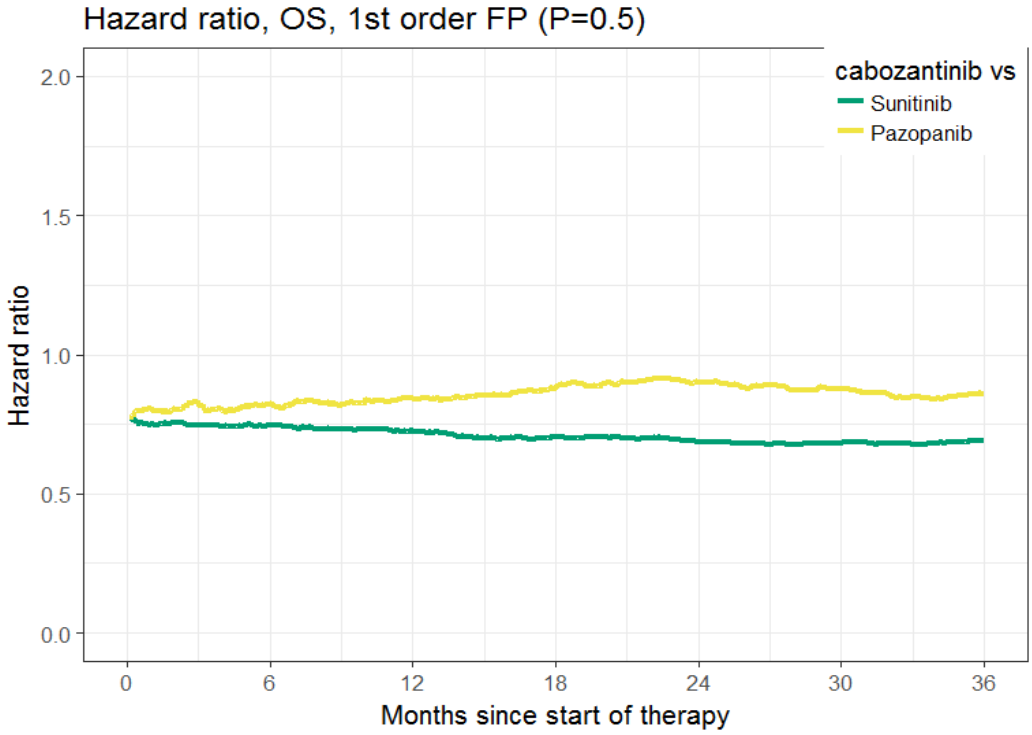


Figure 37 Survival curve, OS; fractional polynomial 1st order (p=-1)

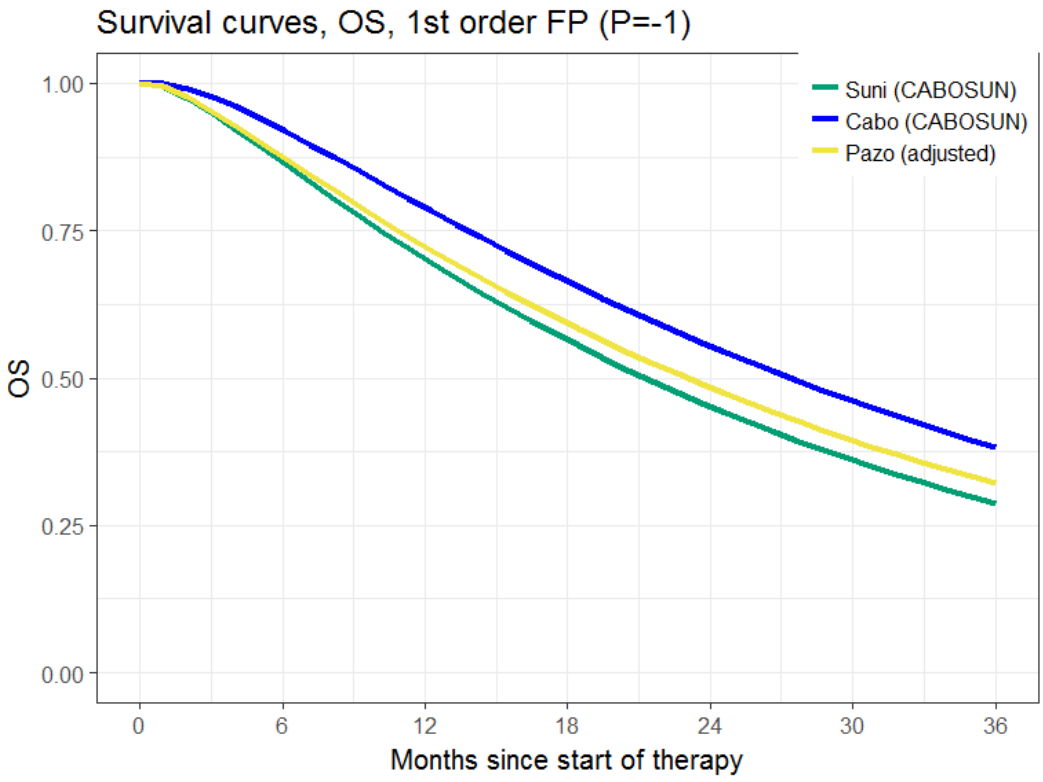


Figure 38 Hazard ratio plot, OS; fractional polynomial 1st order (p=-1)

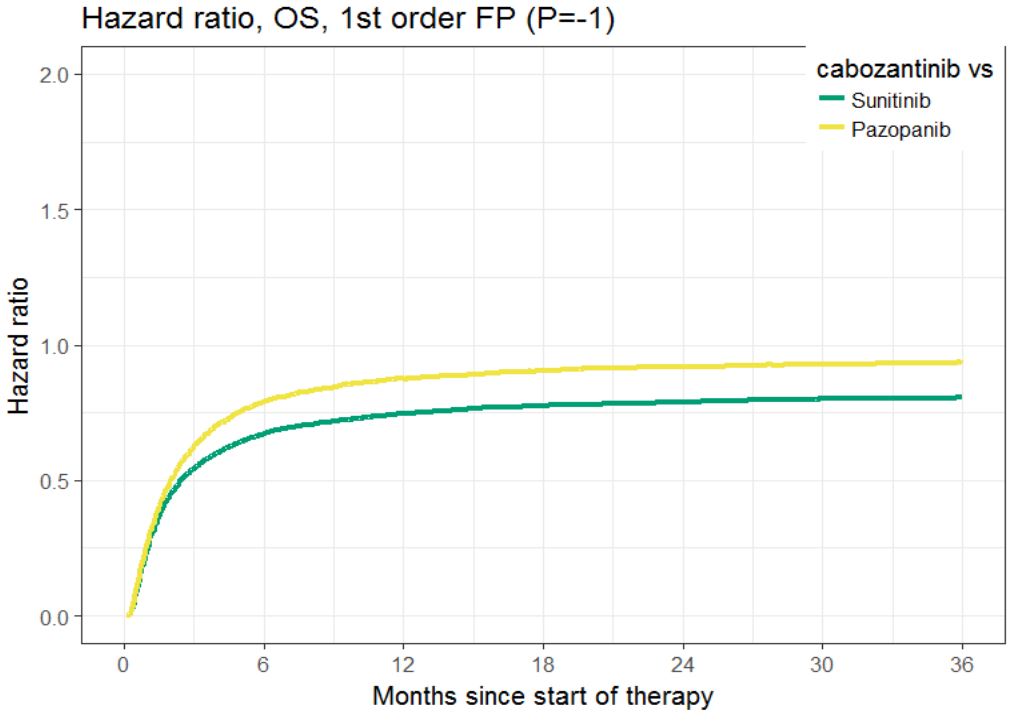


Figure 39 Survival curve, OS; fractional polynomial 1st order (p=-0.5)

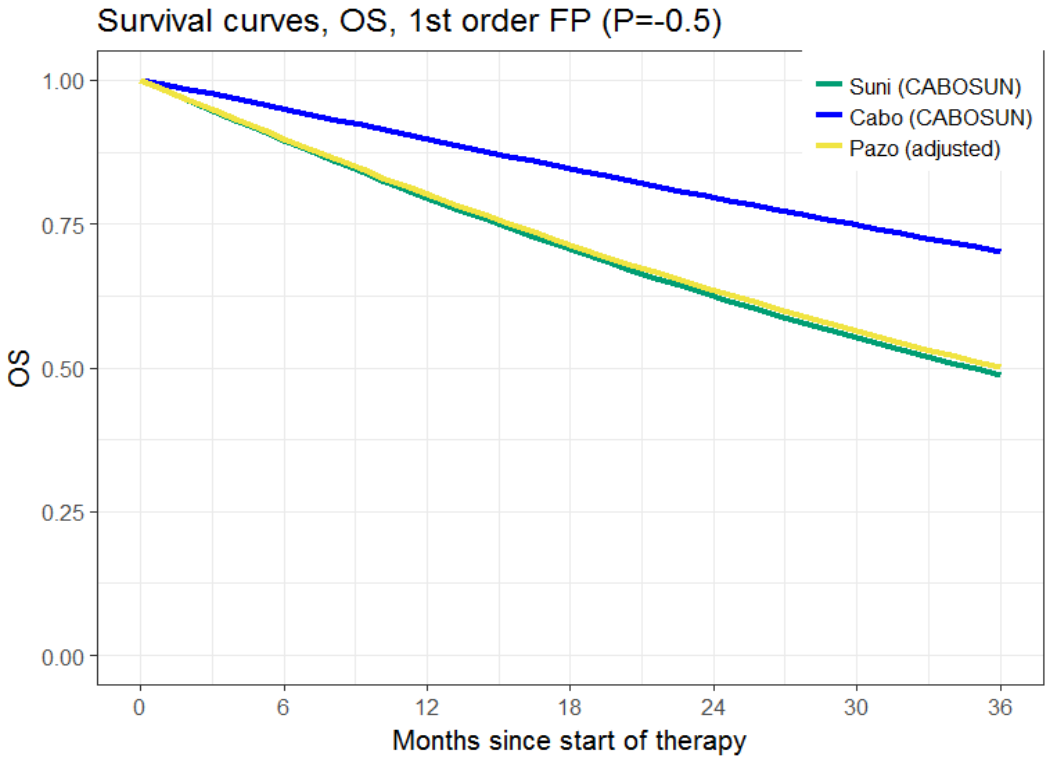


Figure 40 Hazard ratio plot, OS; fractional polynomial 1st order (p=-0.5)

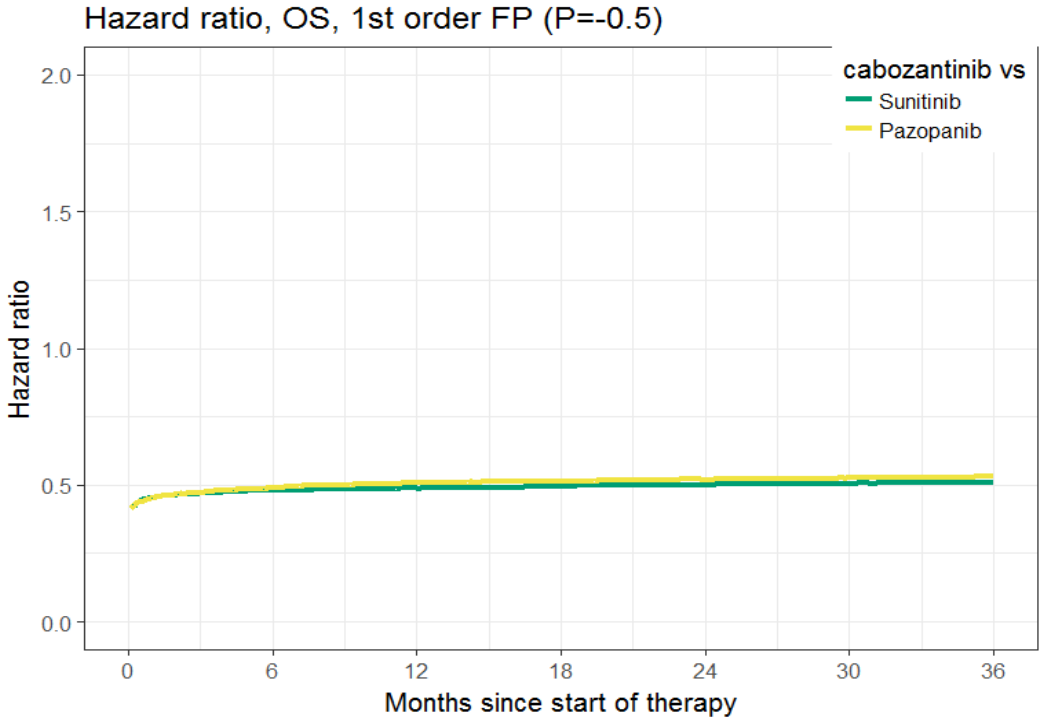


Table 10 Hazard ratios - monthly, OS fractional polynomial 1st order pazopanib

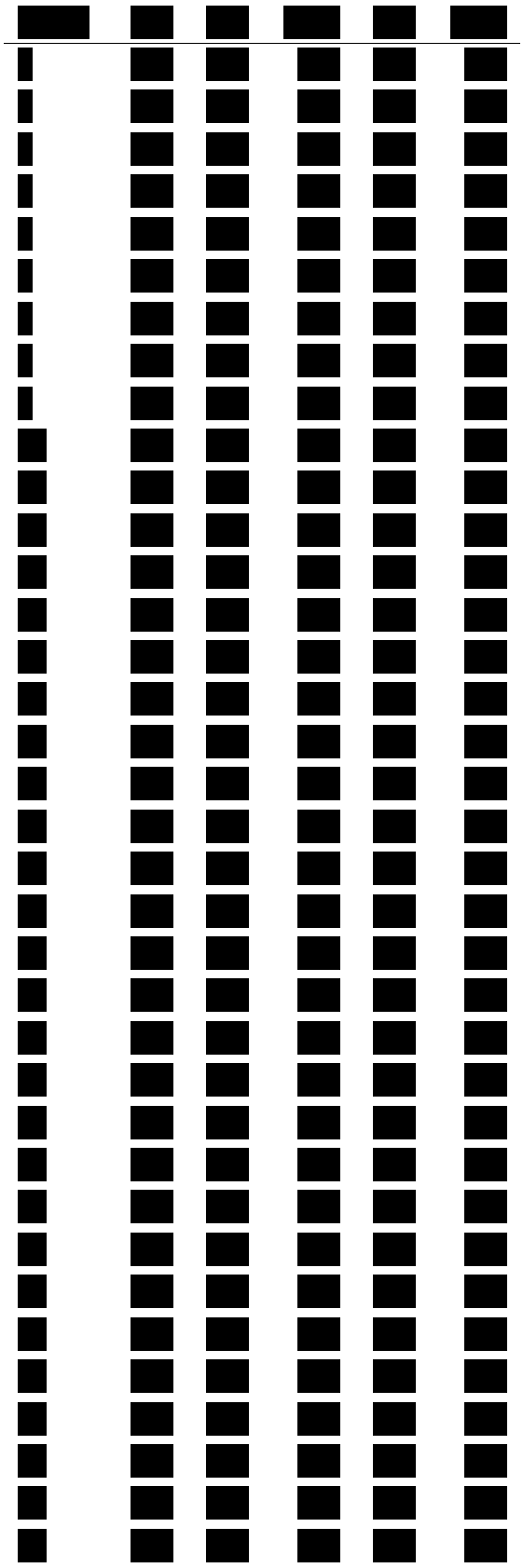
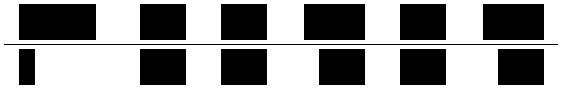
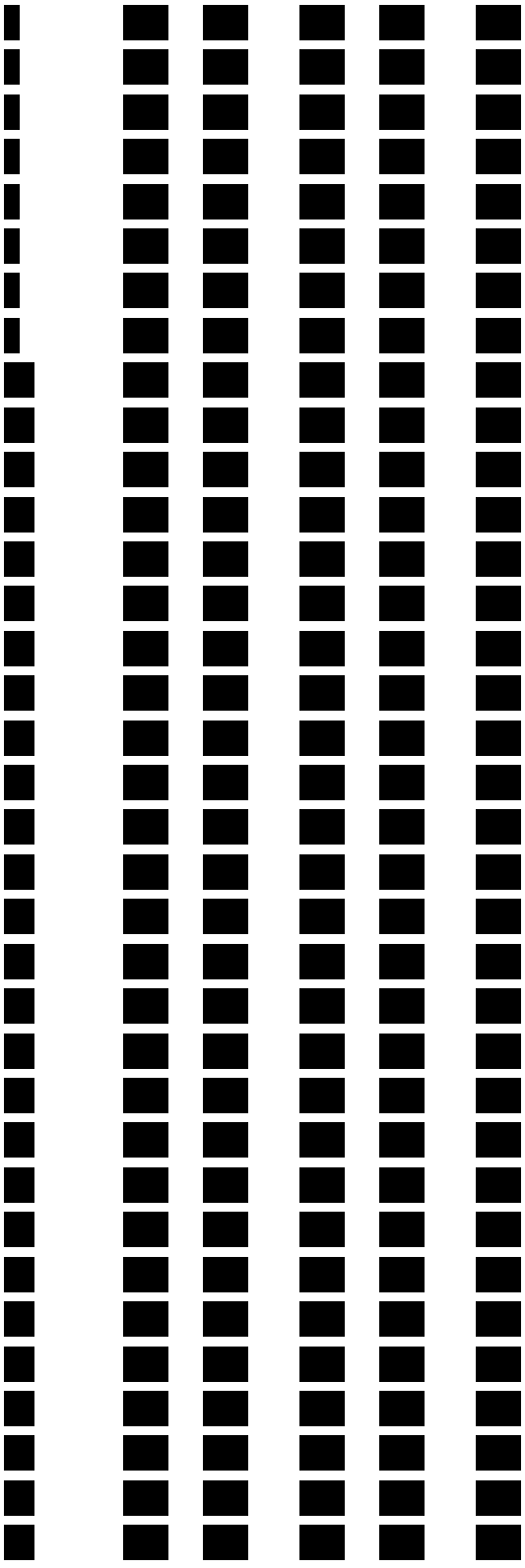


Table 11 Hazard ratios - monthly, OS fractional polynomial 1st order sunitinib





OS fractional polynomial 2nd order

Figure 41 Survival curve, OS; fractional polynomial 2nd order (p1=-0.5, p2=0)

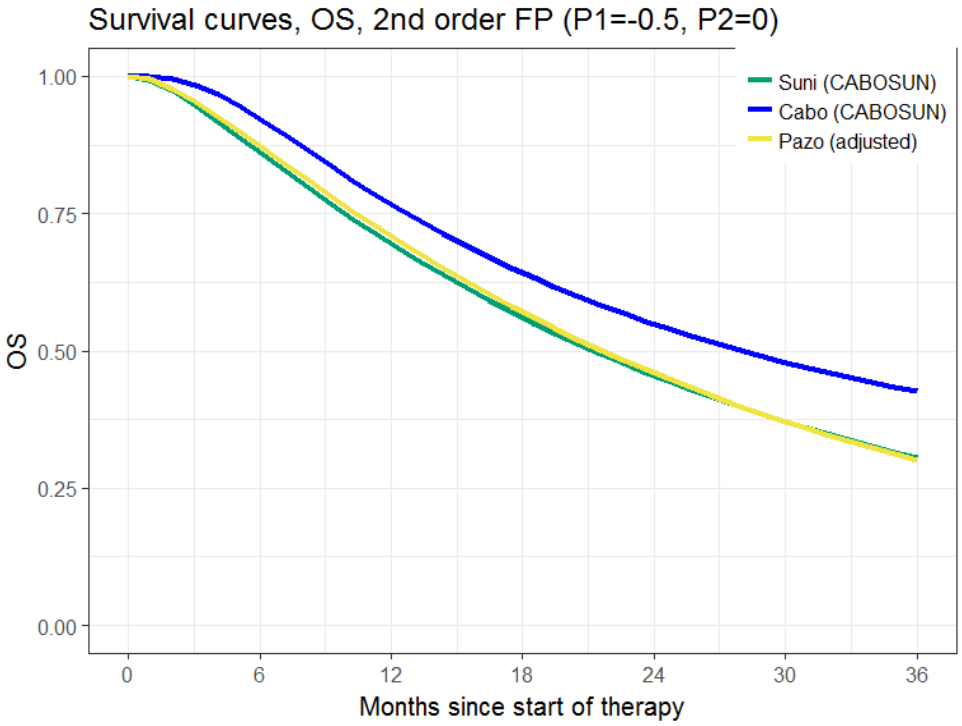


Figure 42 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-0.5, p2=0)

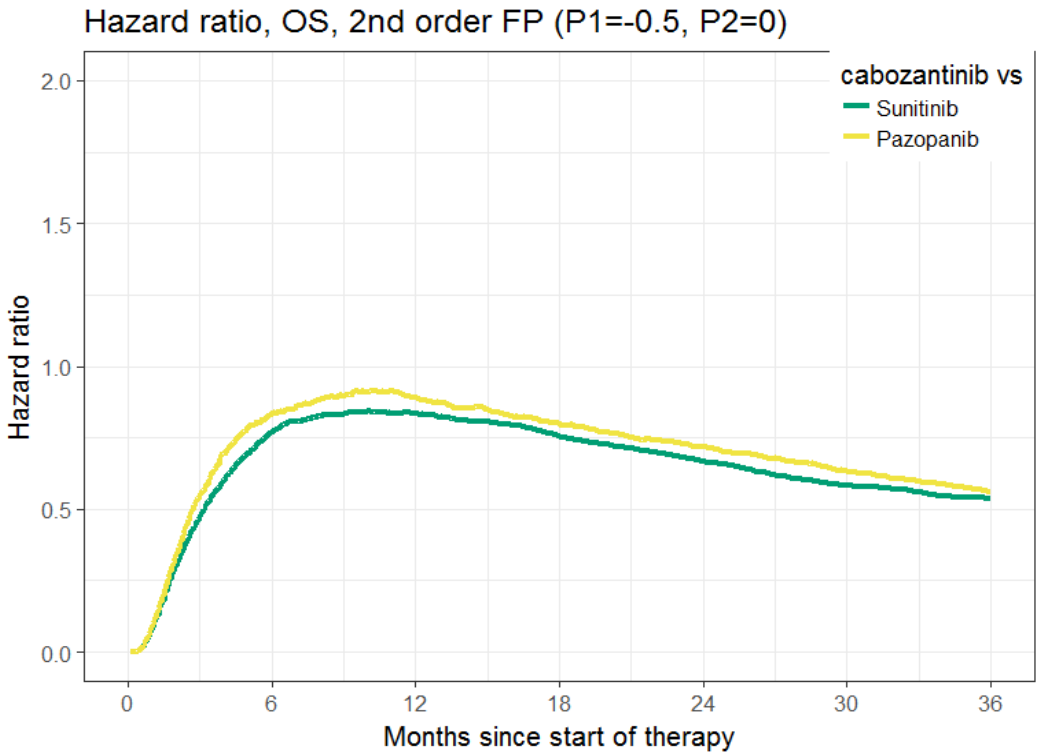


Figure 43 Survival curve, OS; fractional polynomial 2nd order (p1=-1, p2=0)

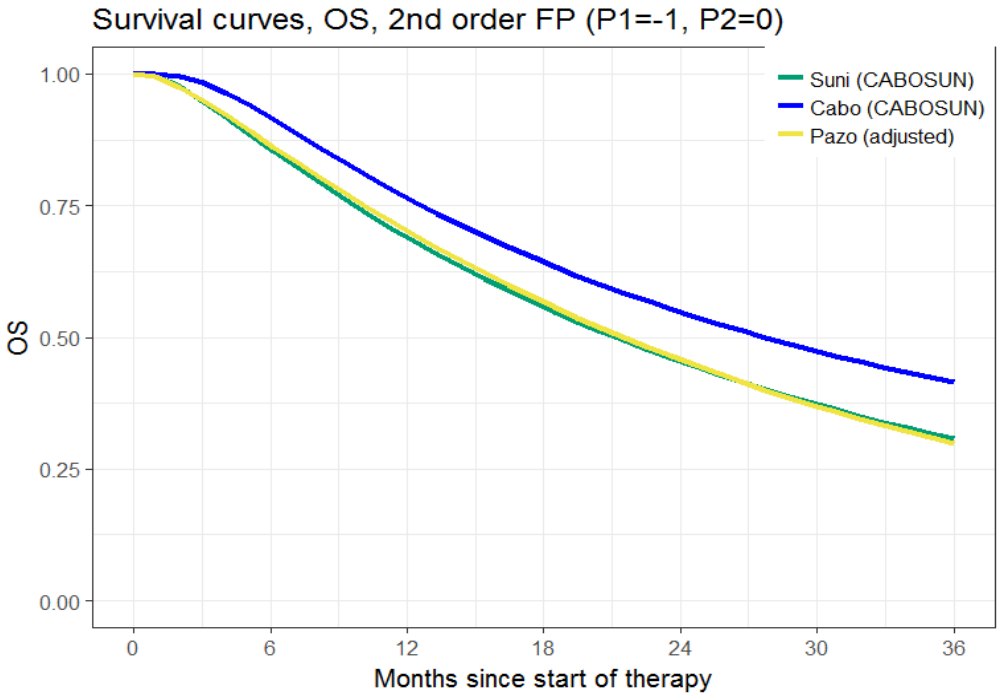


Figure 44 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, p2=0)

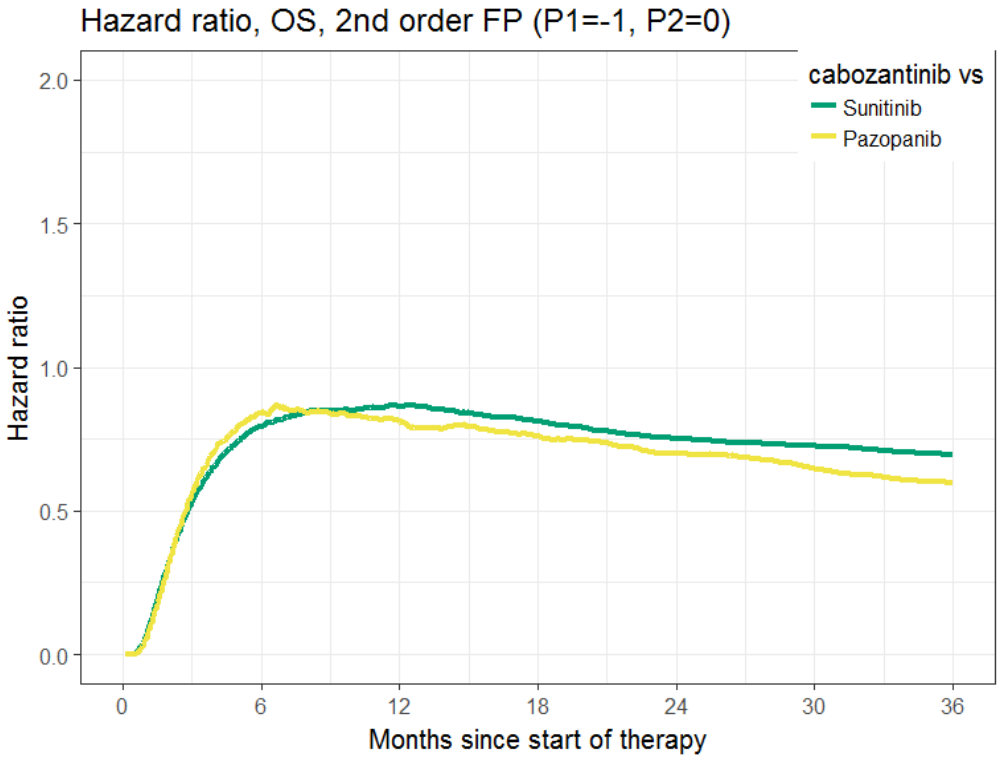


Figure 45 Survival curve, OS; fractional polynomial 2nd order (p1=-1, p2=-1)

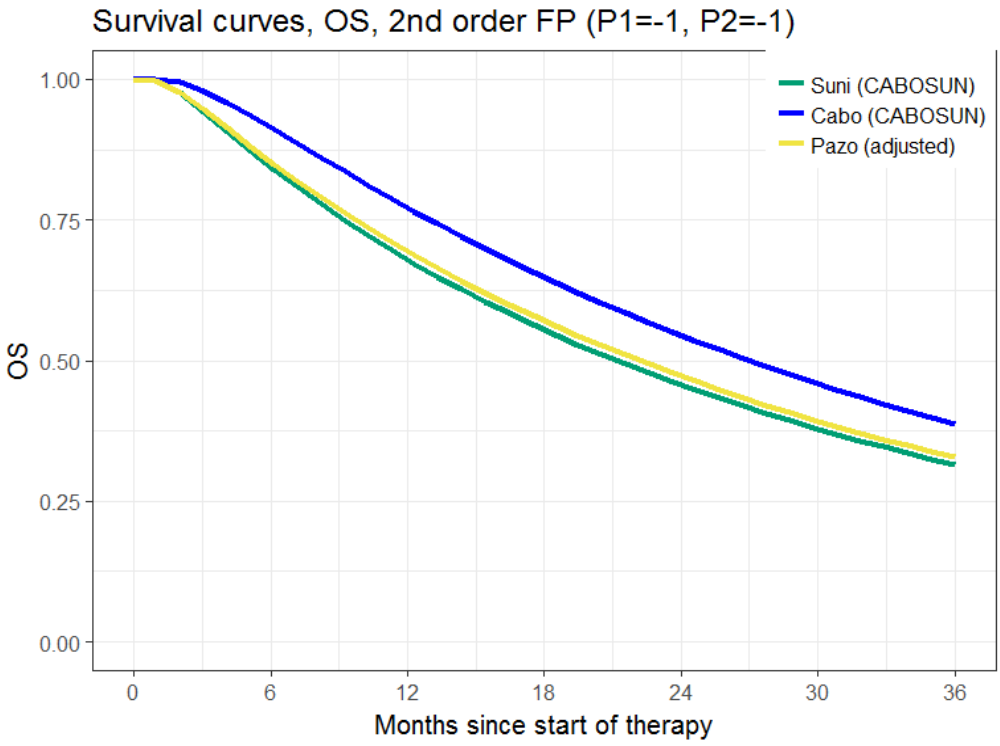


Figure 46 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, p2=-1)

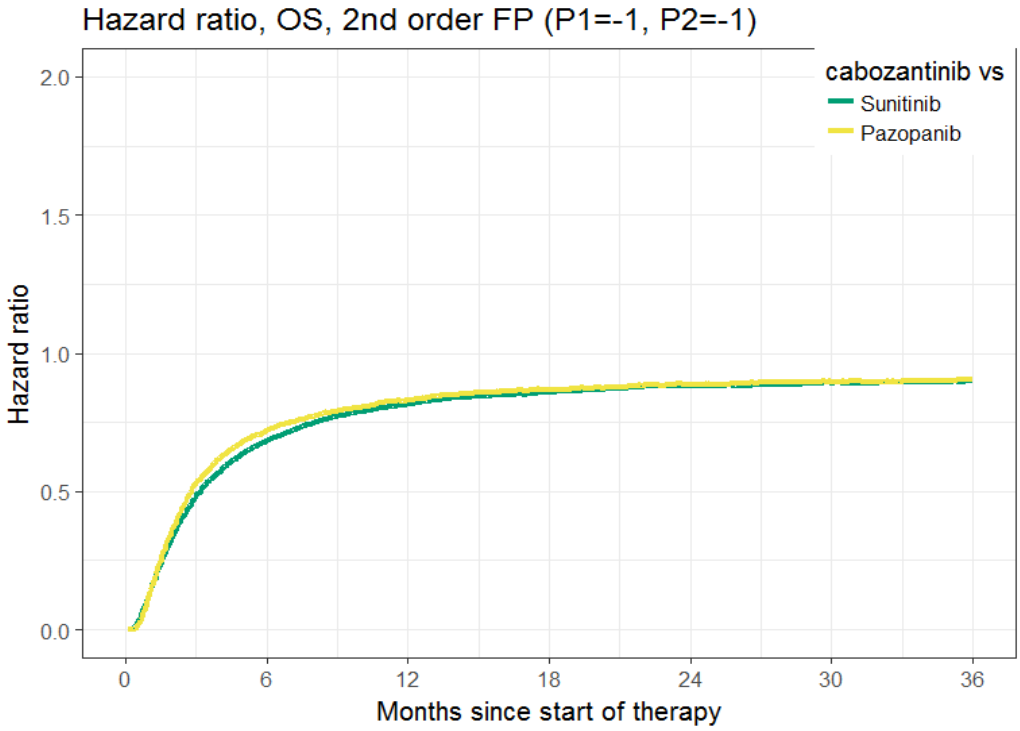


Figure 47 Survival curve, OS; fractional polynomial 2nd order (p1=-1, p2=0.5)

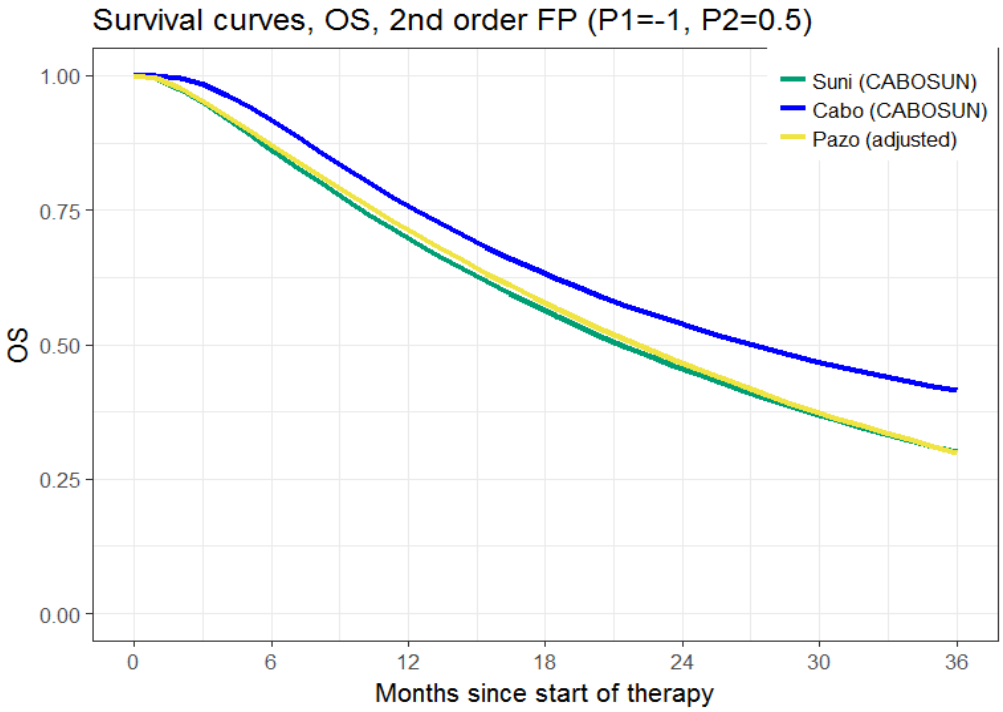


Figure 48 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, p2=0.5)

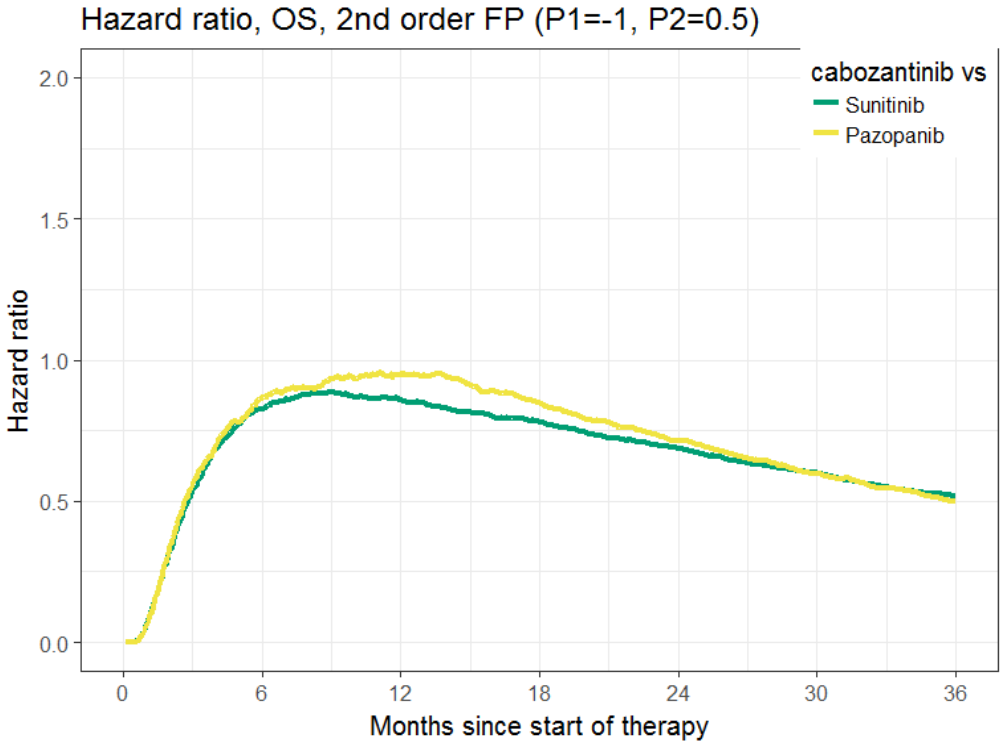


Figure 49 Survival curve plot, OS; fractional polynomial 2nd order (p1=-1, P2=1)

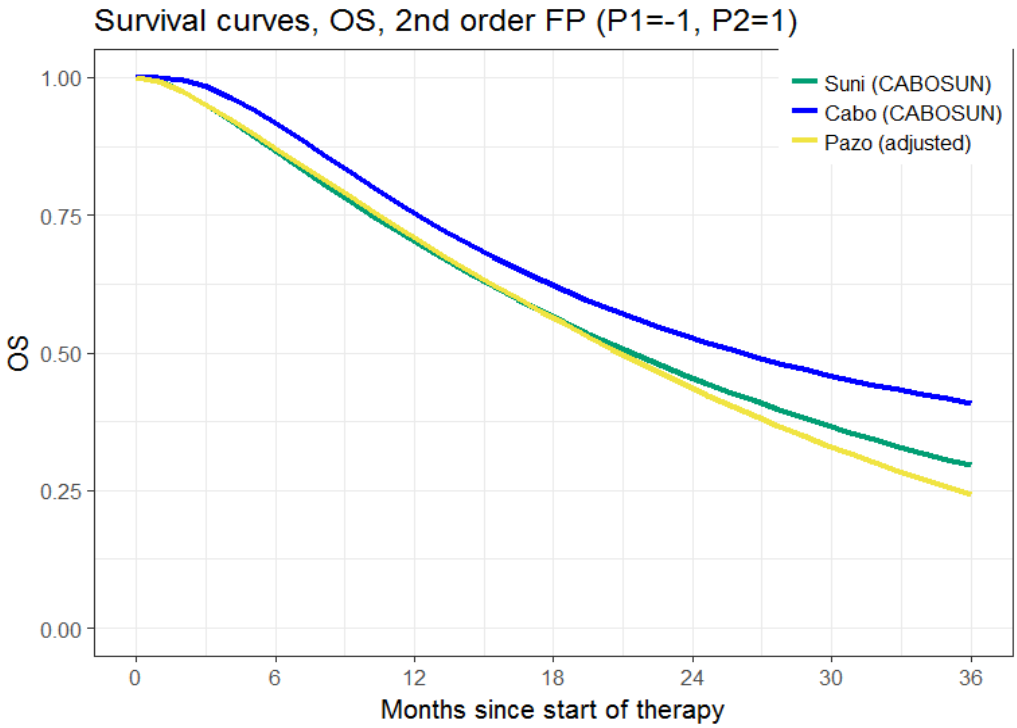


Figure 50 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, P2=1)

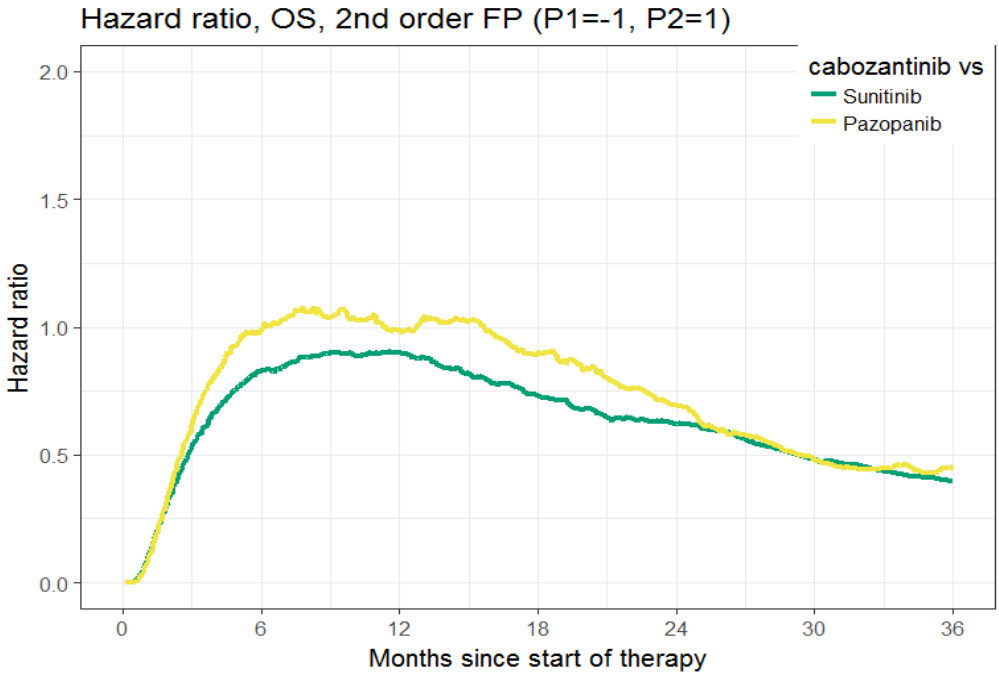


Table 12 Hazard ratios - monthly, OS fractional polynomial 2nd order pazopanib

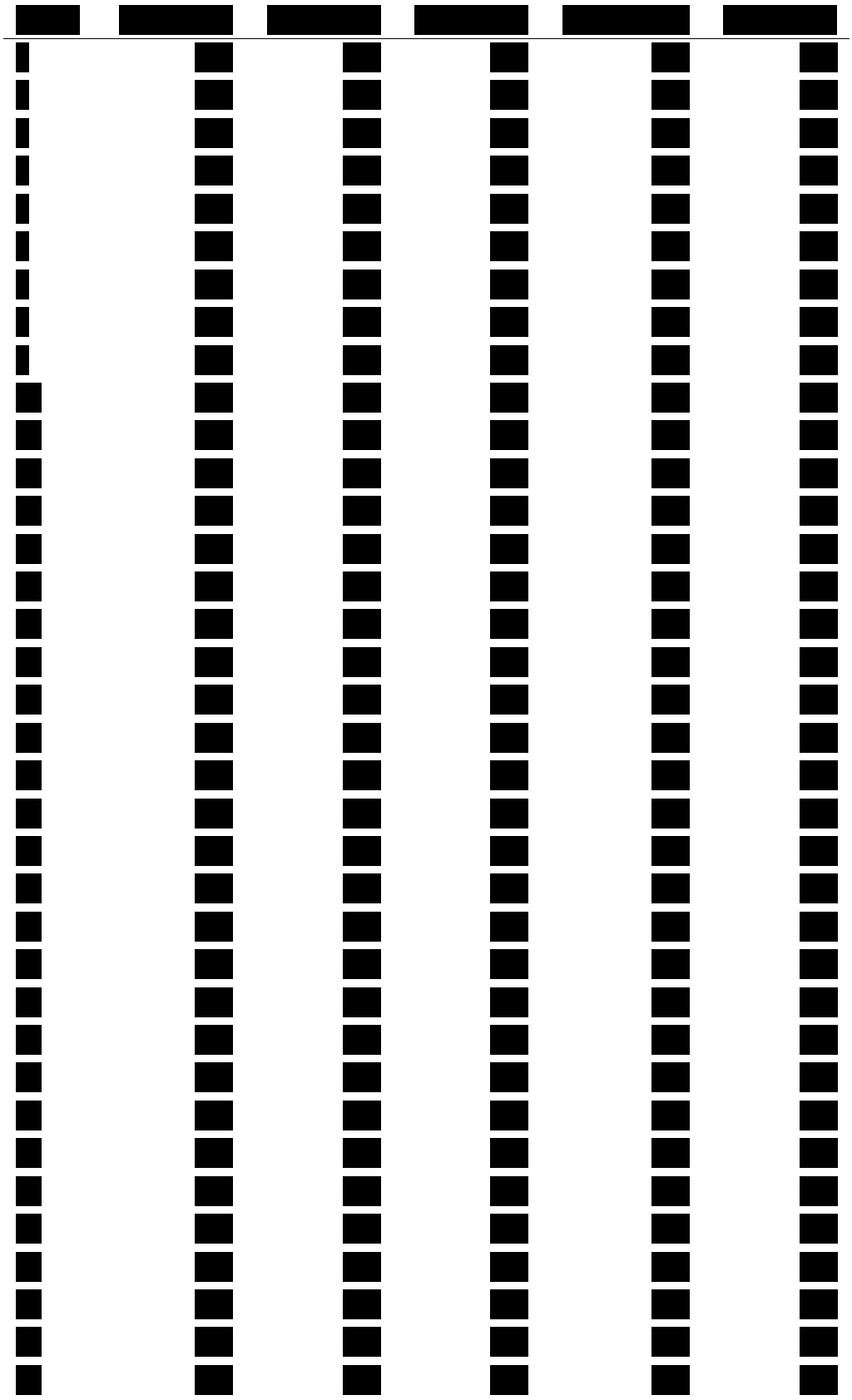
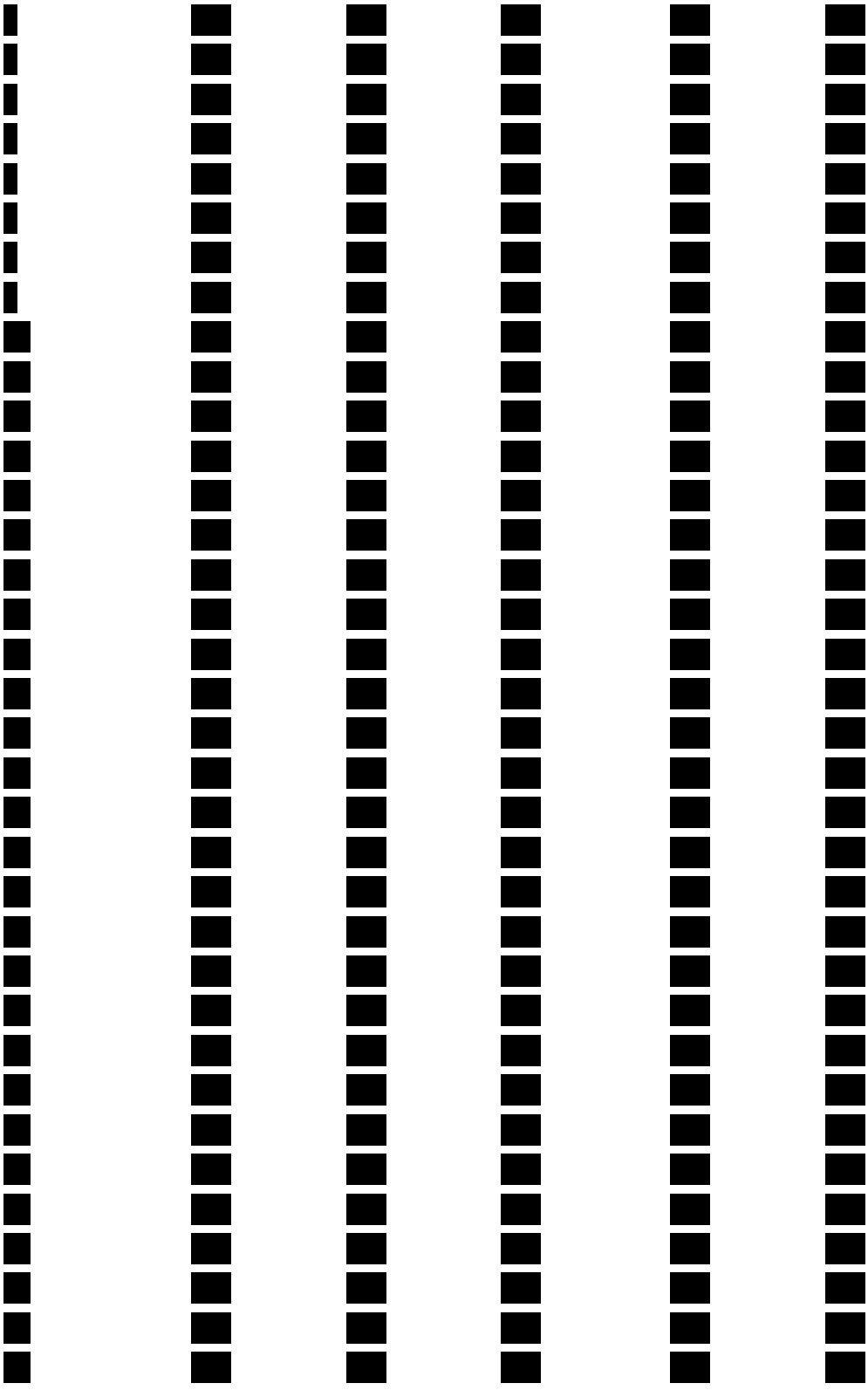


Table 13 Hazard ratios - monthly, OS fractional polynomial 2nd order sunitinib



PFS fractional polynomial 1st order

Figure 51 Survival curve, PFS; fractional polynomial 1st order (p=0)

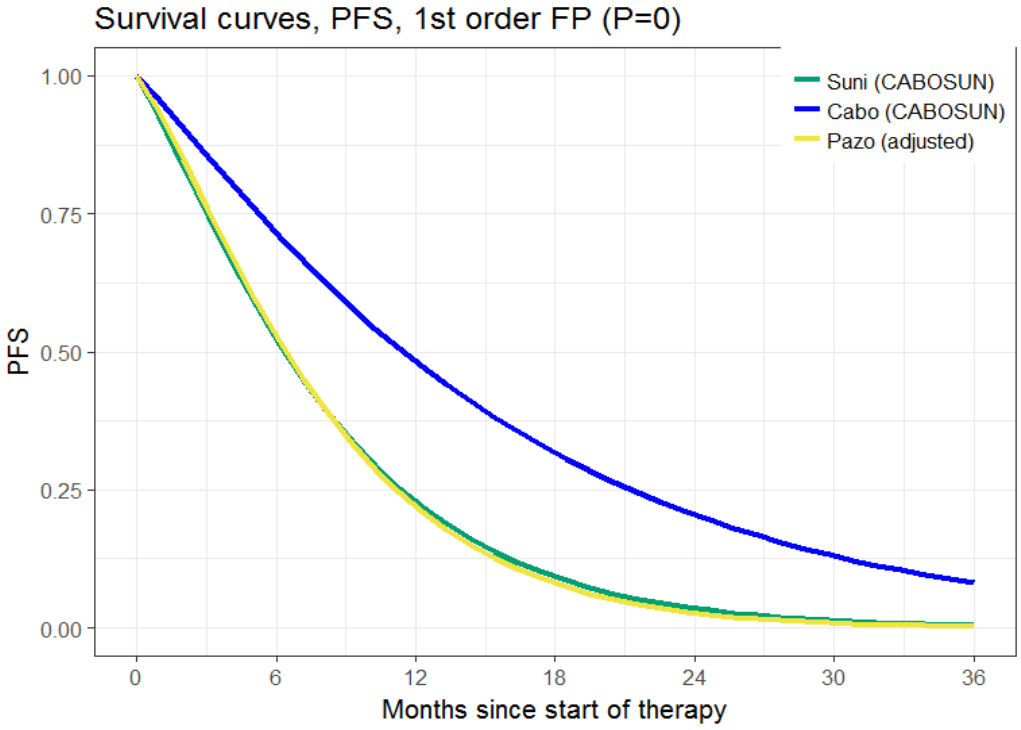


Figure 52 Hazard ratio plot, PFS; fractional polynomial 1st order (p=0)

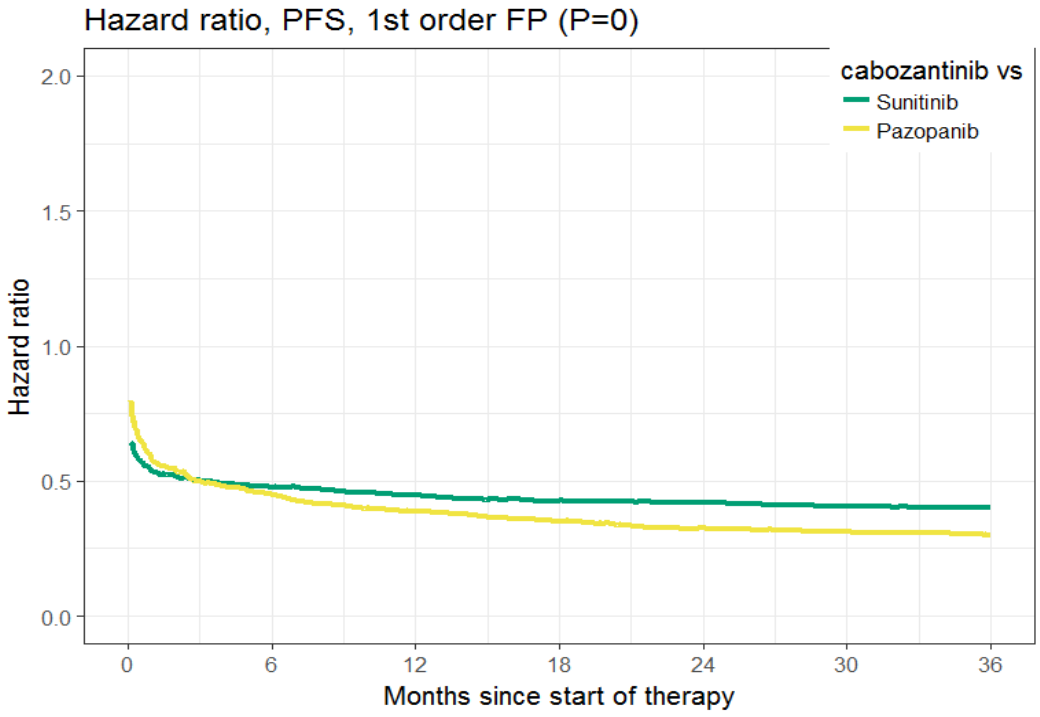


Figure 53 Survival curve, PFS; fractional polynomial 1st order (p=1)

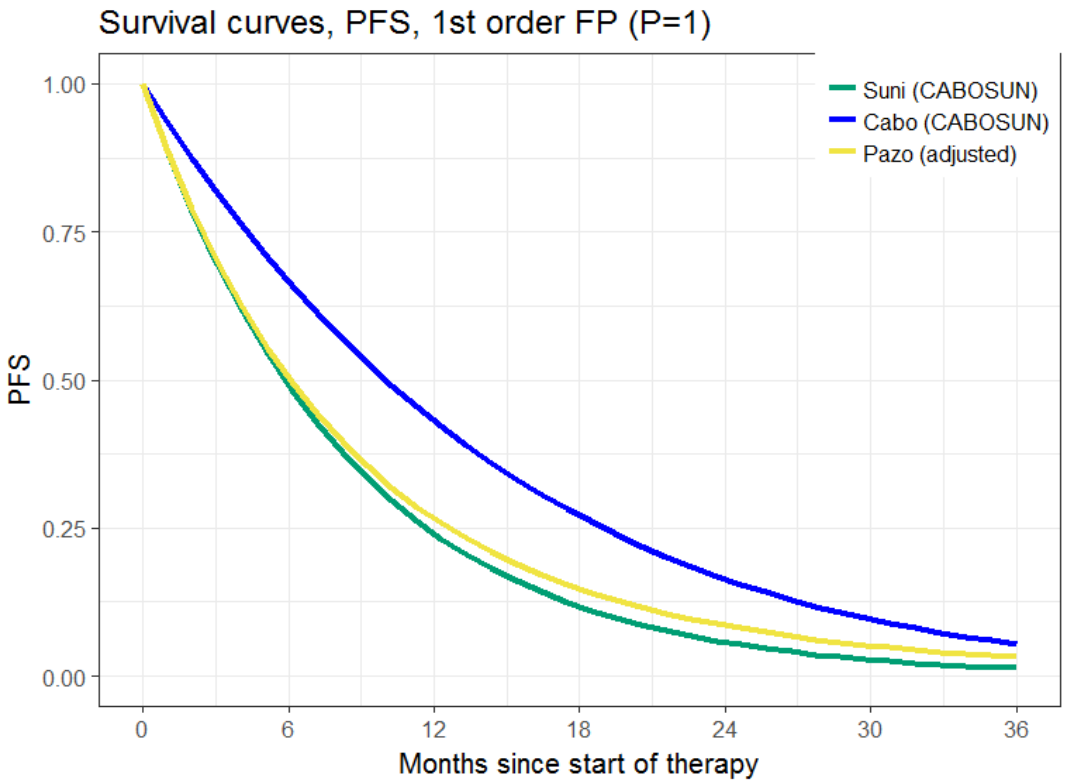


Figure 54 Hazard ratio plot, PFS; fractional polynomial 1st order (p=1)

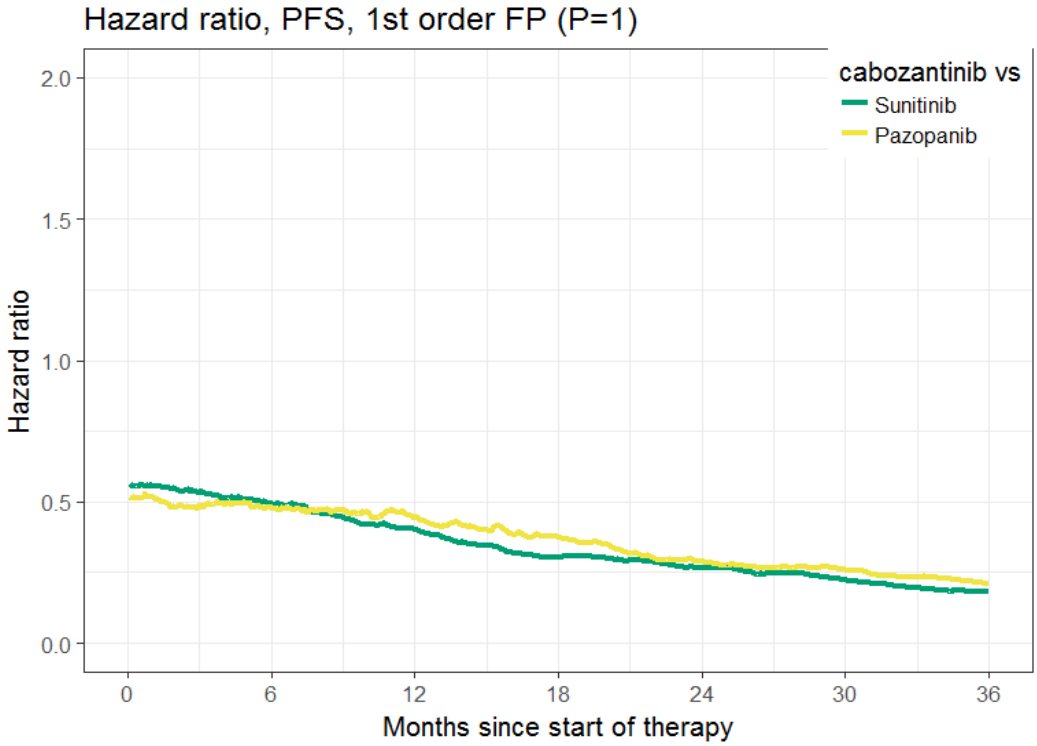


Figure 55 Survival curve, PFS; fractional polynomial 1st order (p=0.5)

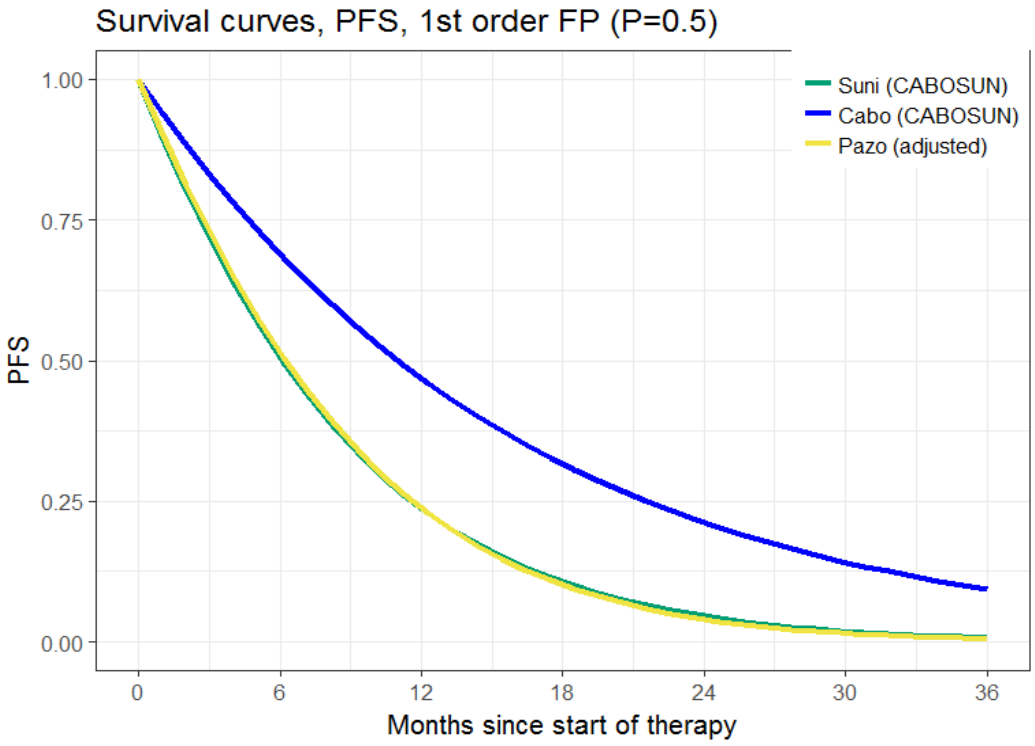


Figure 56 Hazard ratio plot, PFS; fractional polynomial 1st order (p=0.5)

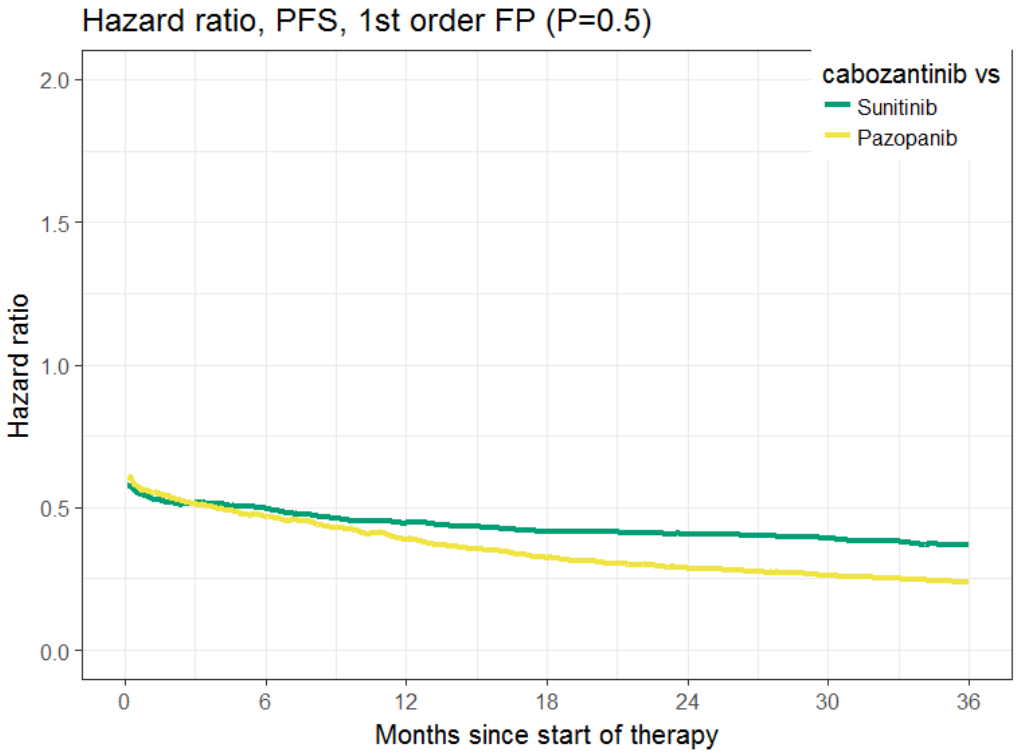


Figure 57 Survival curve, PFS; fractional polynomial 1st order (p=-1)

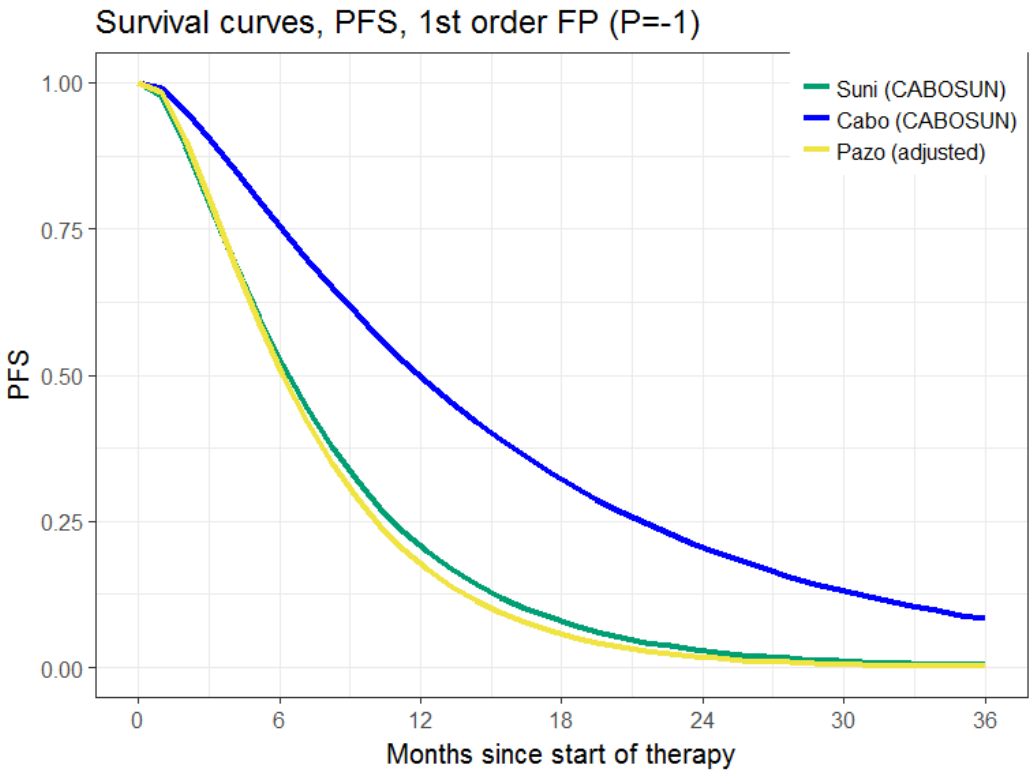


Figure 58 Hazard ratio plot, PFS; fractional polynomial 1st order (p=-1)

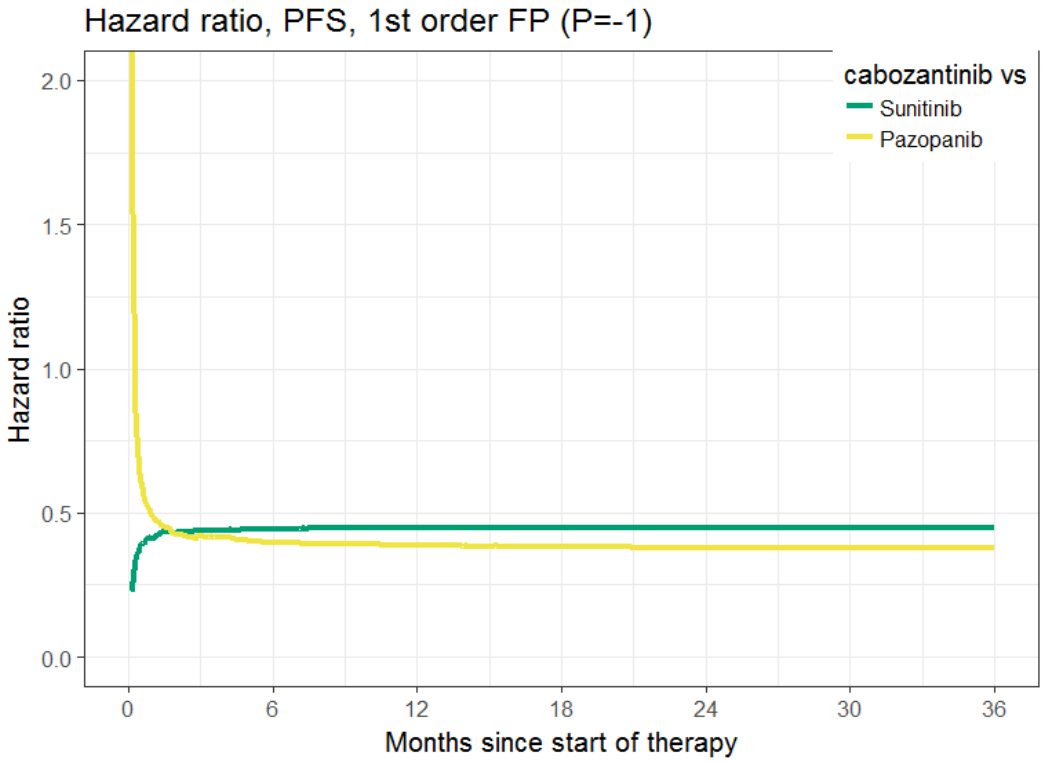


Figure 59 Survival curve, PFS; fractional polynomial 1st order (p=-0.5)

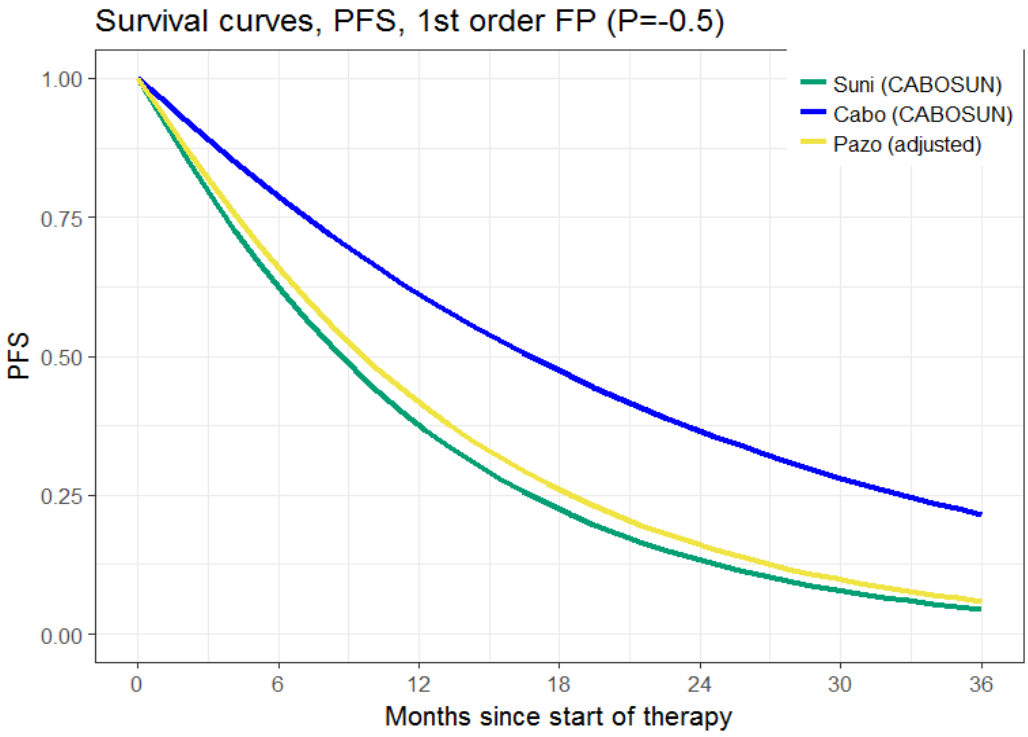


Figure 60 Hazard ratio plot, PFS; fractional polynomial 1st order (p=-0.5)

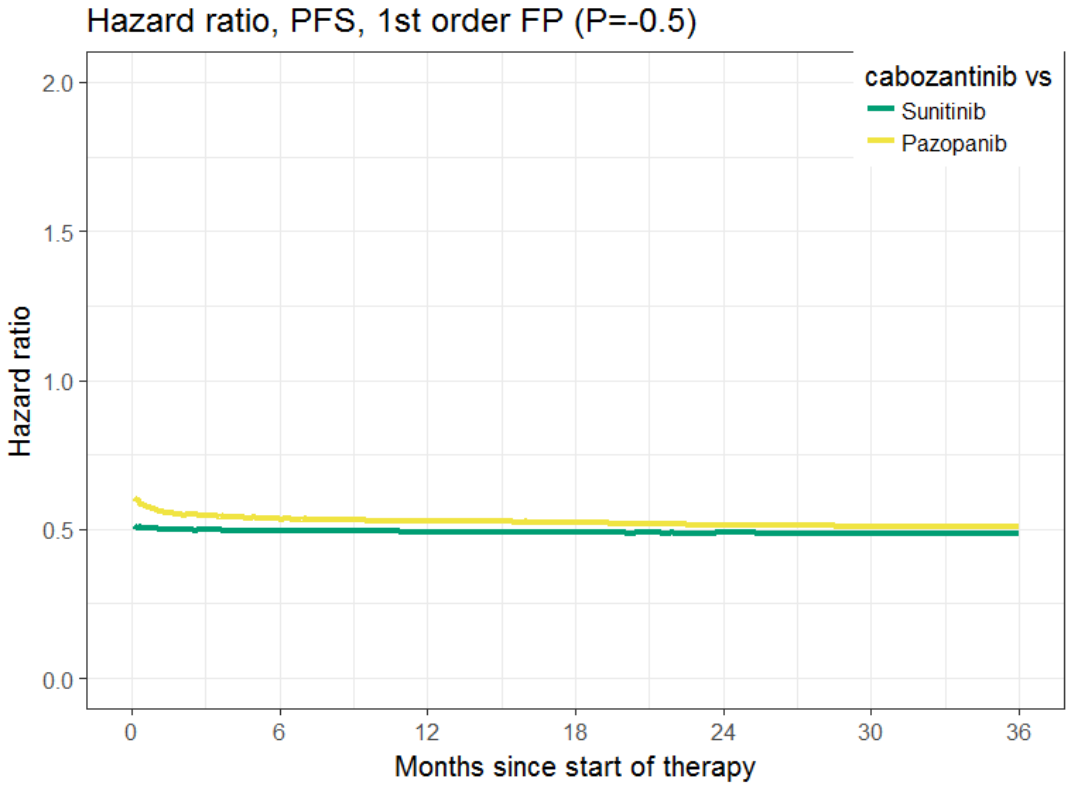


Table 14 Hazard ratios - monthly, PFS fractional polynomial 1st order pazopanib

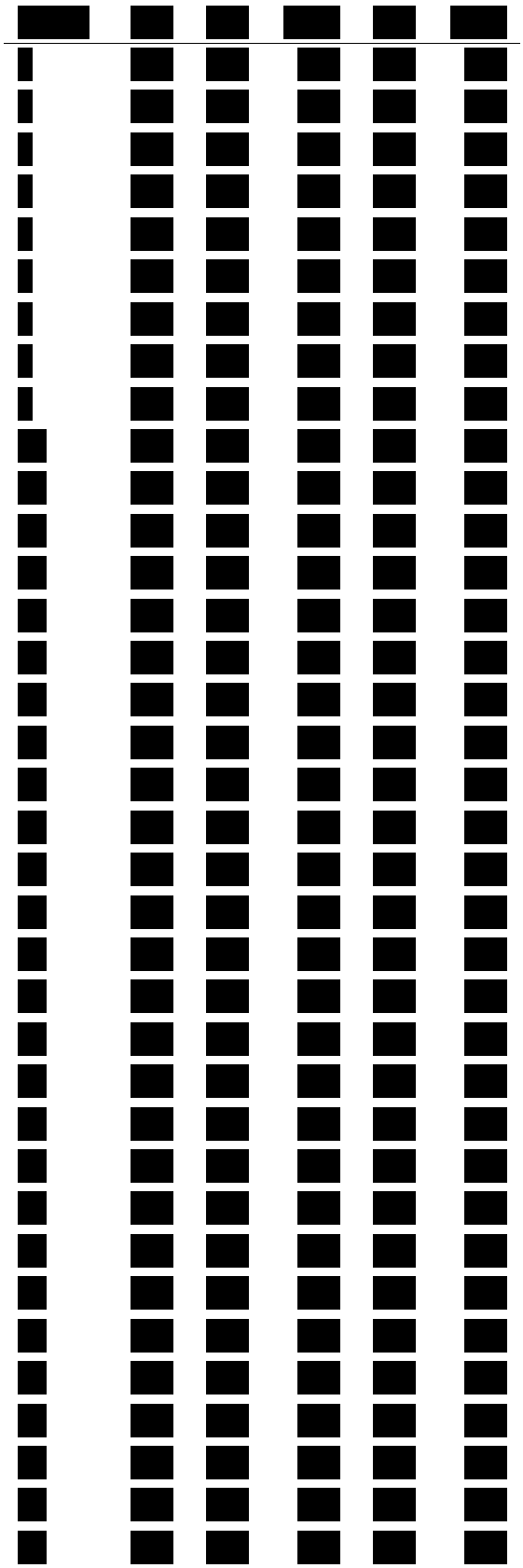
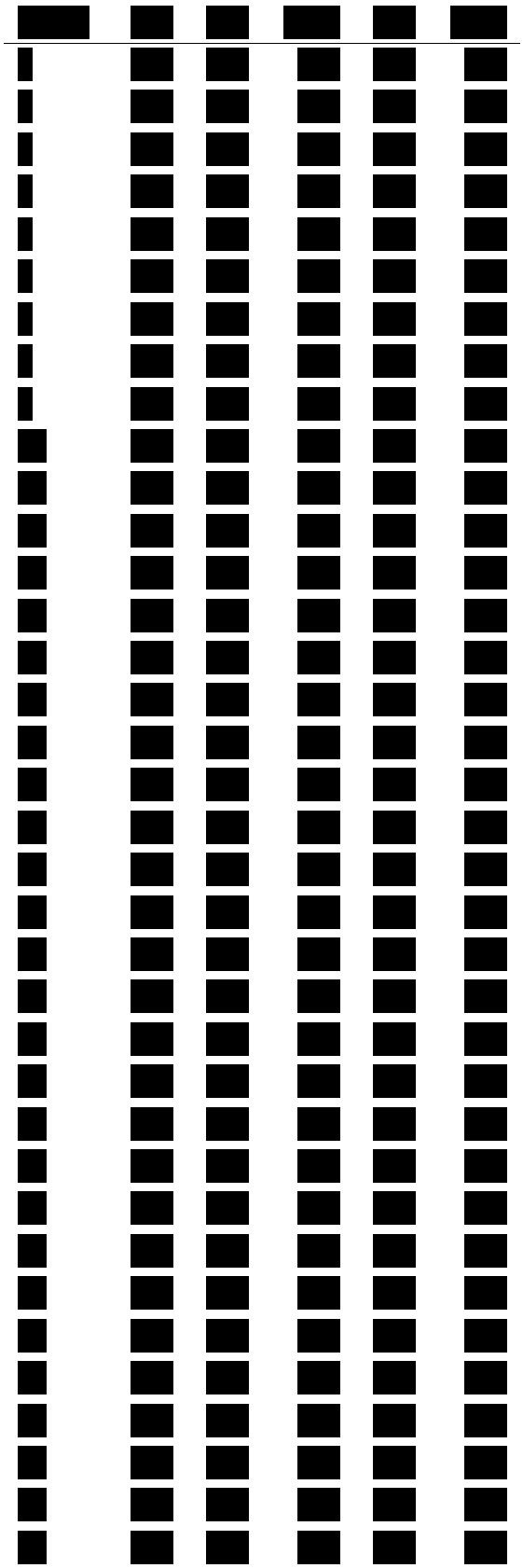


Table 15 Hazard ratios - monthly, PFS fractional polynomial 1st order sunitinib



PFS fractional polynomial 2nd order

Figure 61 Survival curve, PFS; fractional polynomial 2nd order (p1=-0.5, p2=0)

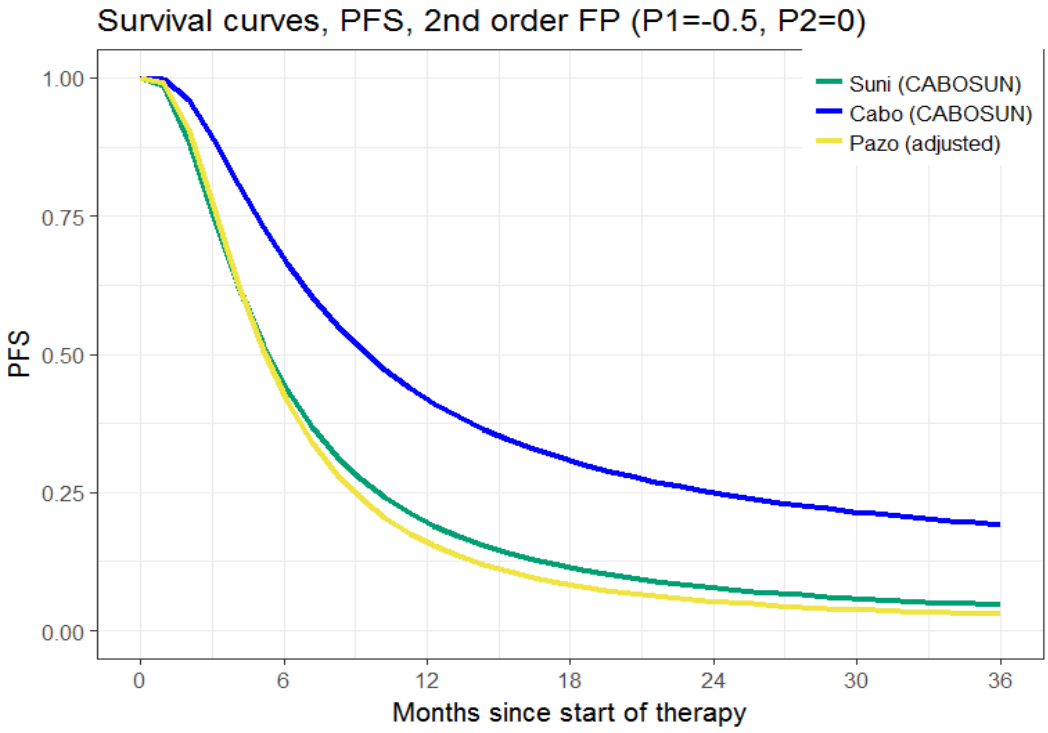


Figure 62 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-0.5, p2=0)

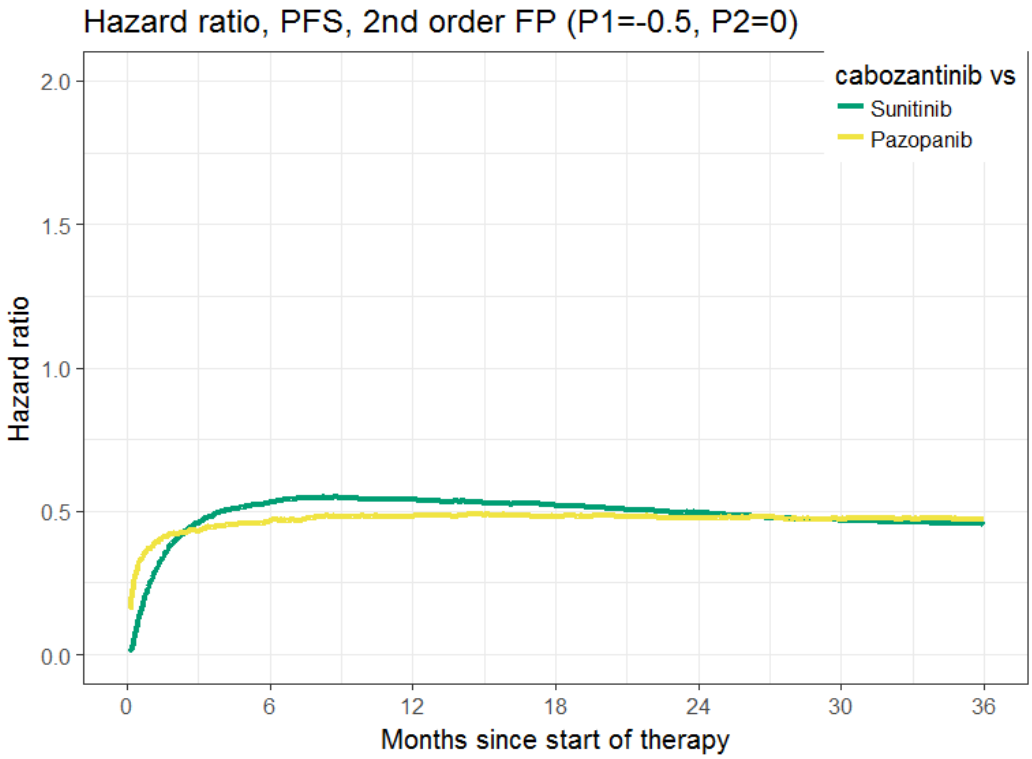


Figure 63 Survival curve, PFS; fractional polynomial 2nd order (p1=-1, p2=0)

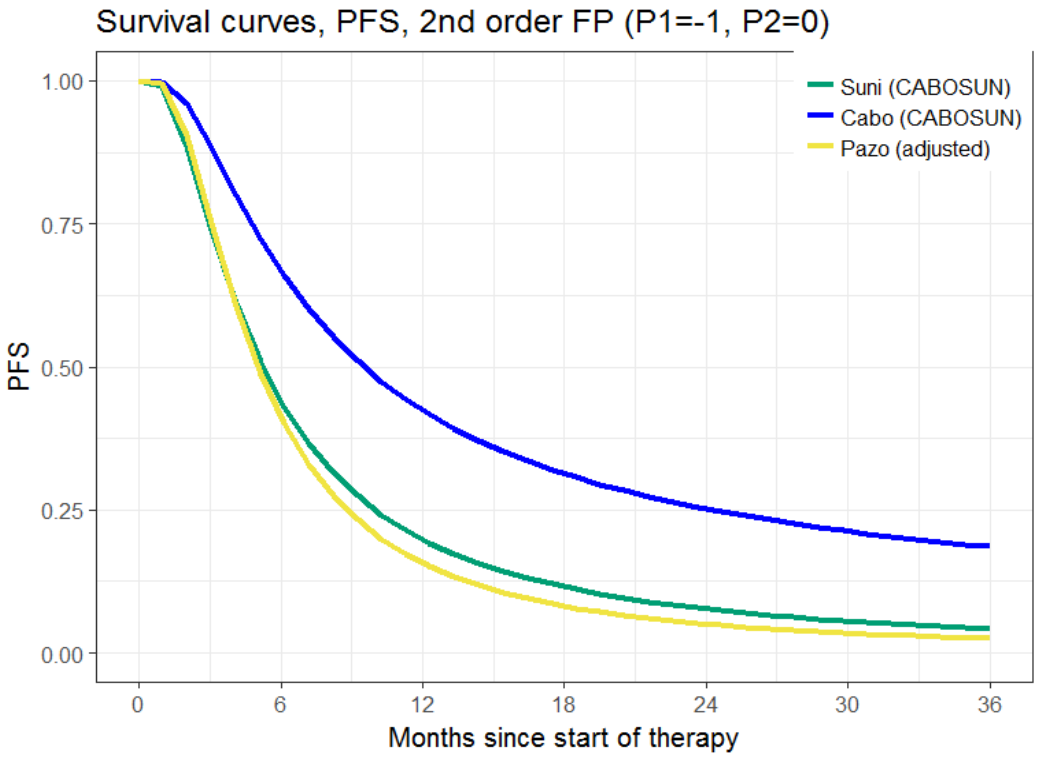


Figure 64 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=0)

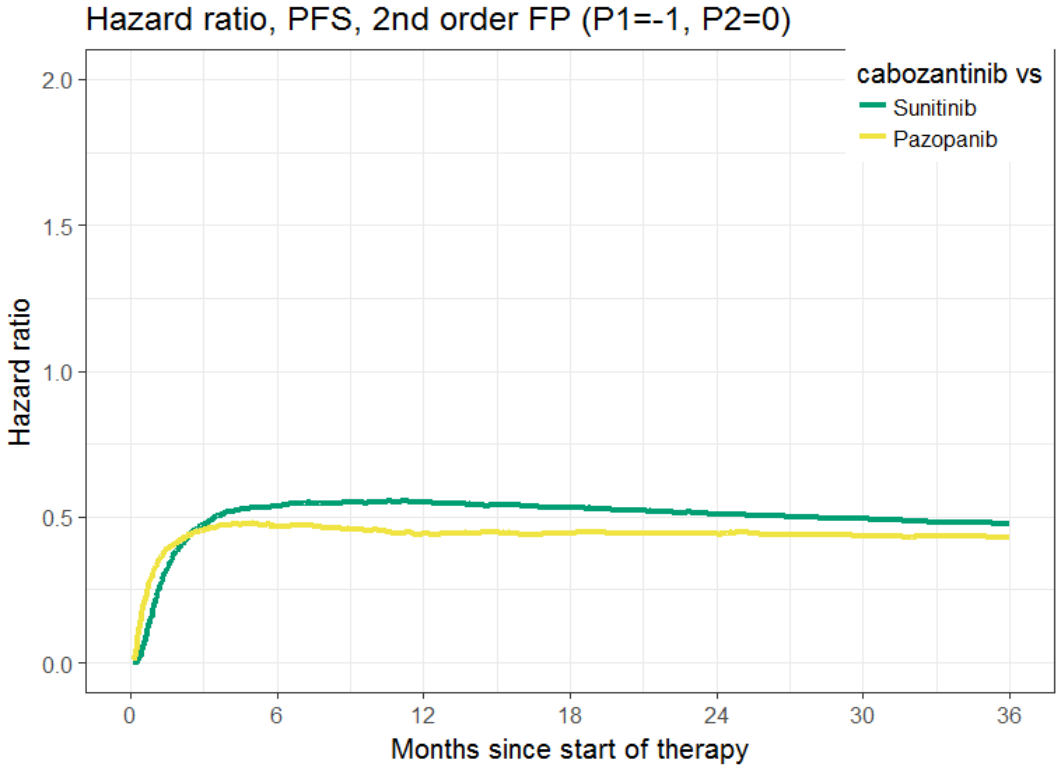


Figure 65 Survival curve, PFS; fractional polynomial 2nd order (p1=-1, p2=-1)

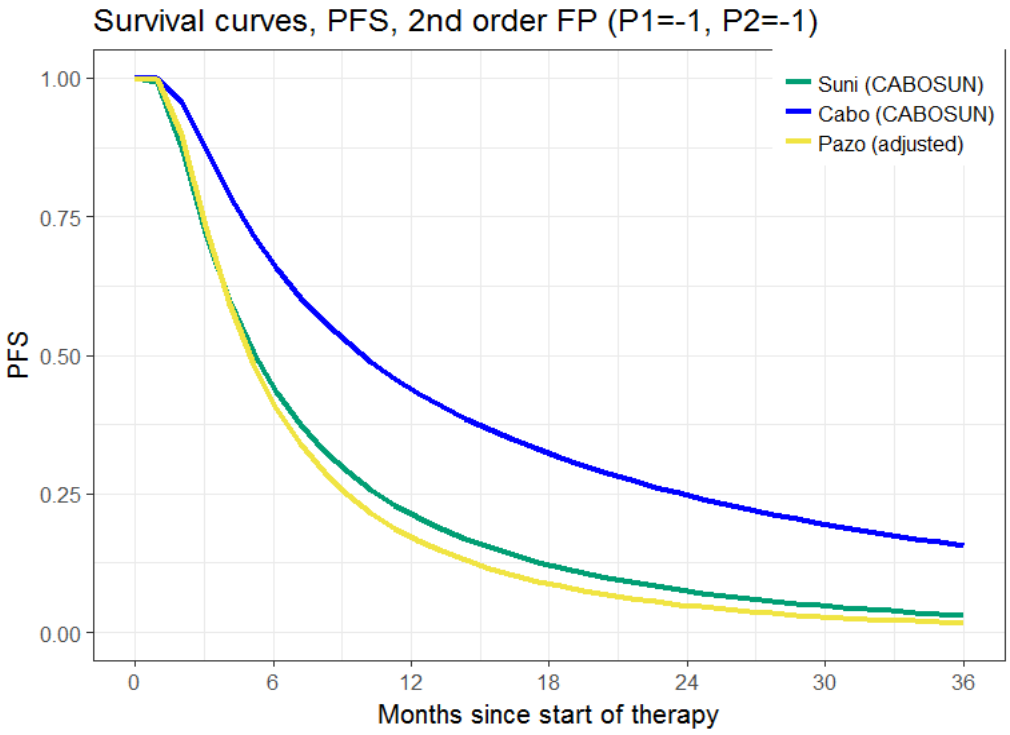


Figure 66 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=-1)

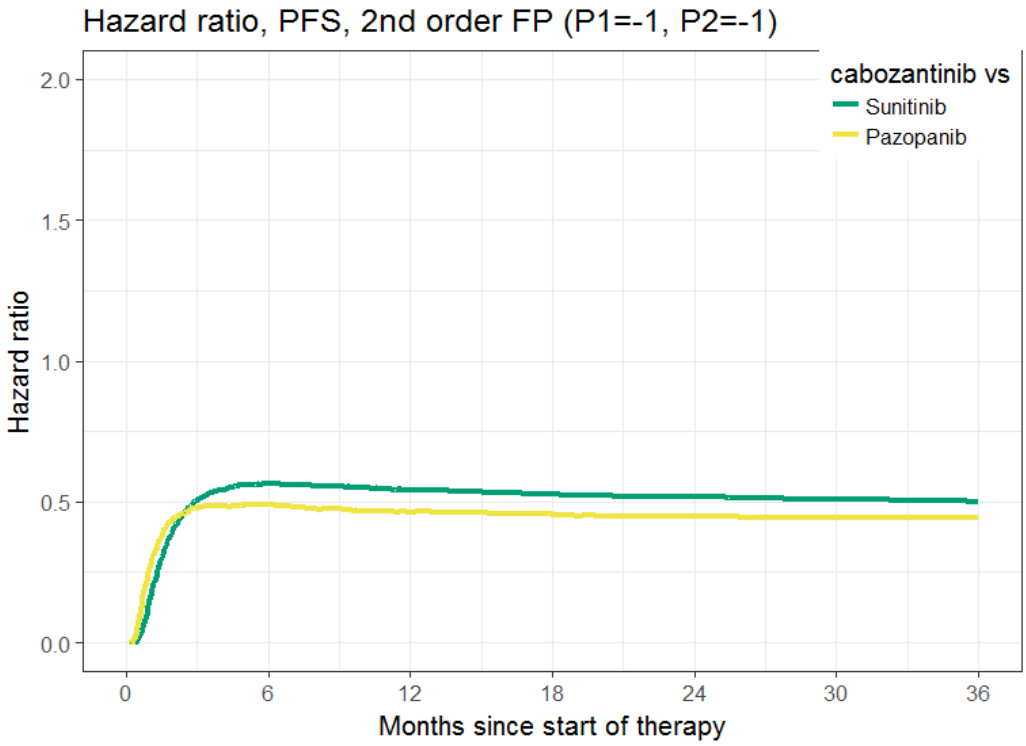


Figure 67 Survival curve, PFS; fractional polynomial 2nd order (p1=-1, p2=0.5)

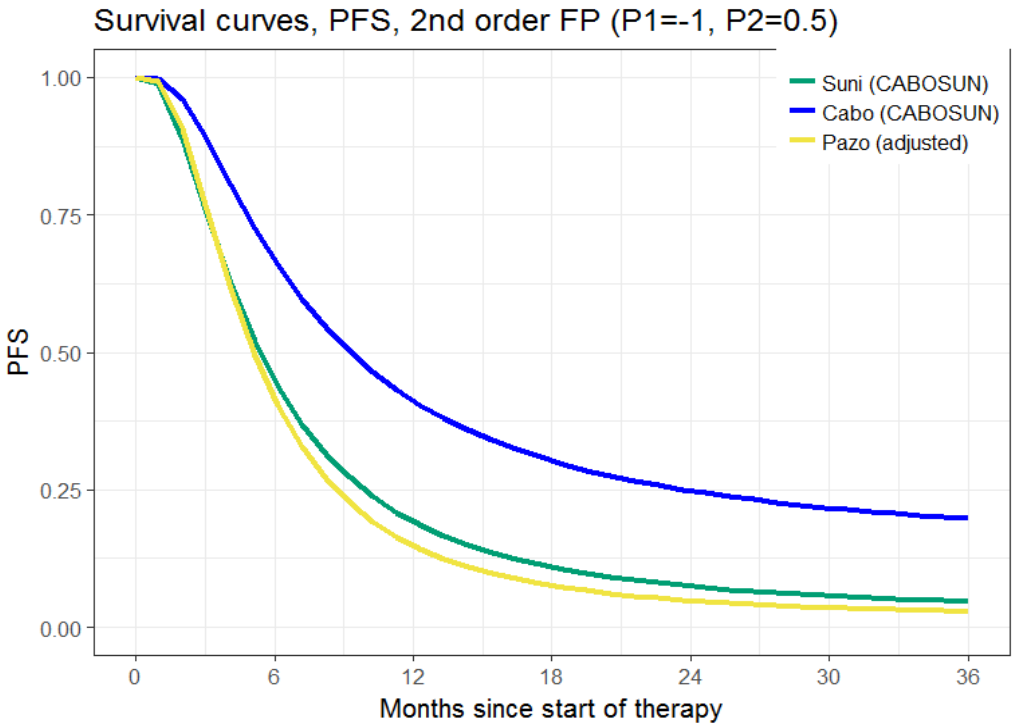


Figure 68 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=0.5)

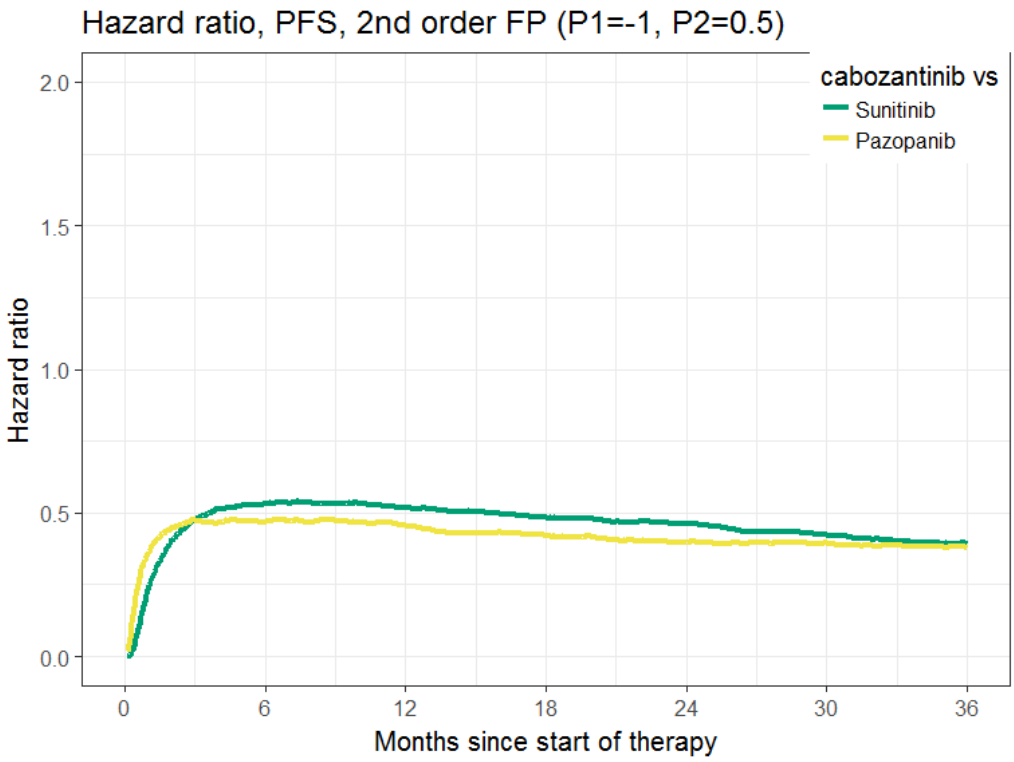


Figure 69 Survival curve, PFS; fractional polynomial 2nd order (p1=-1, p2=1)

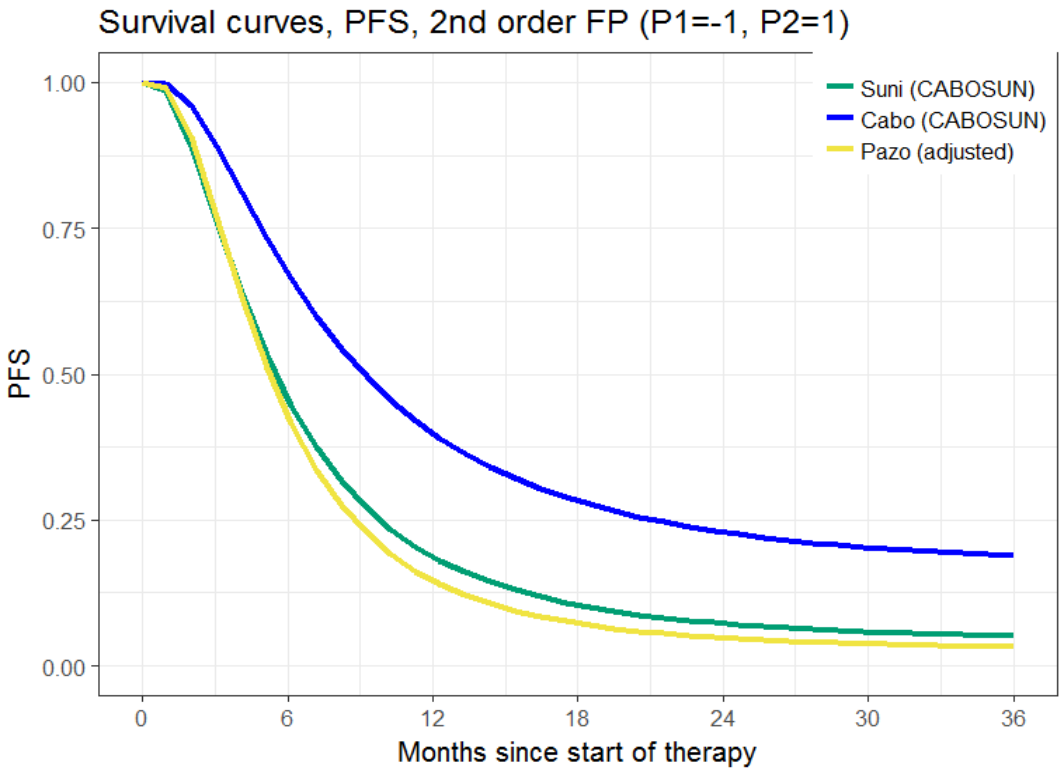


Figure 70 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=1)

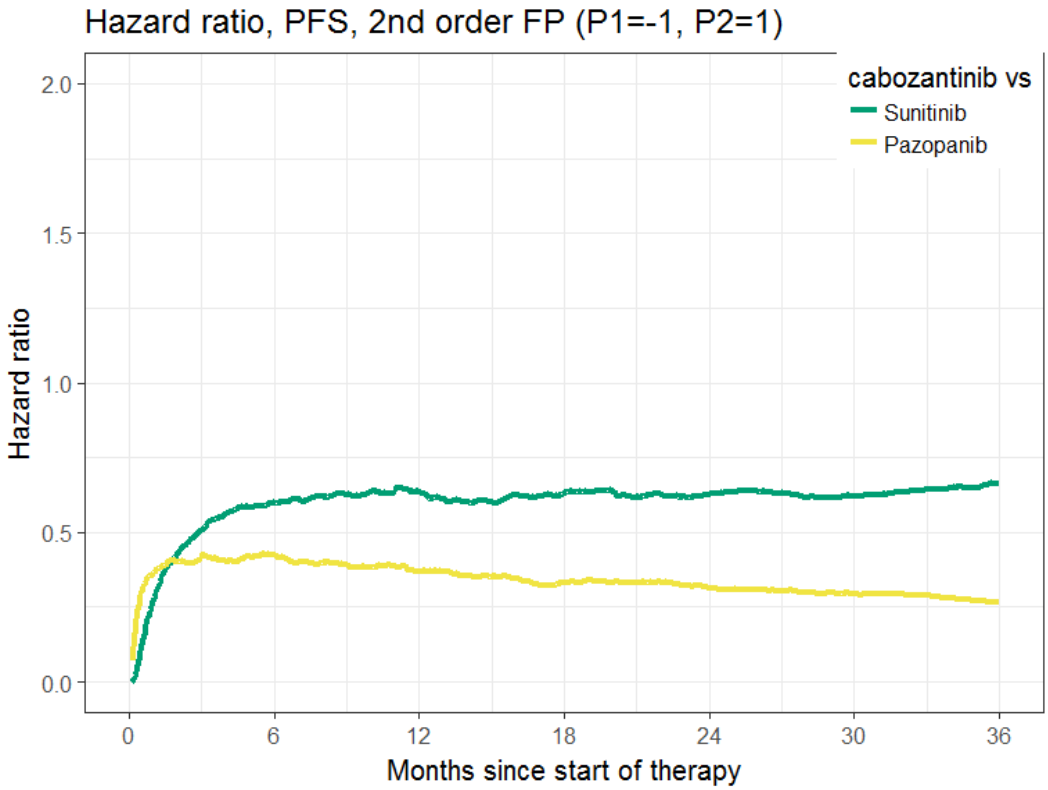


Table 16 Hazard ratios - monthly, OS fractional polynomial 2nd order pazopanib

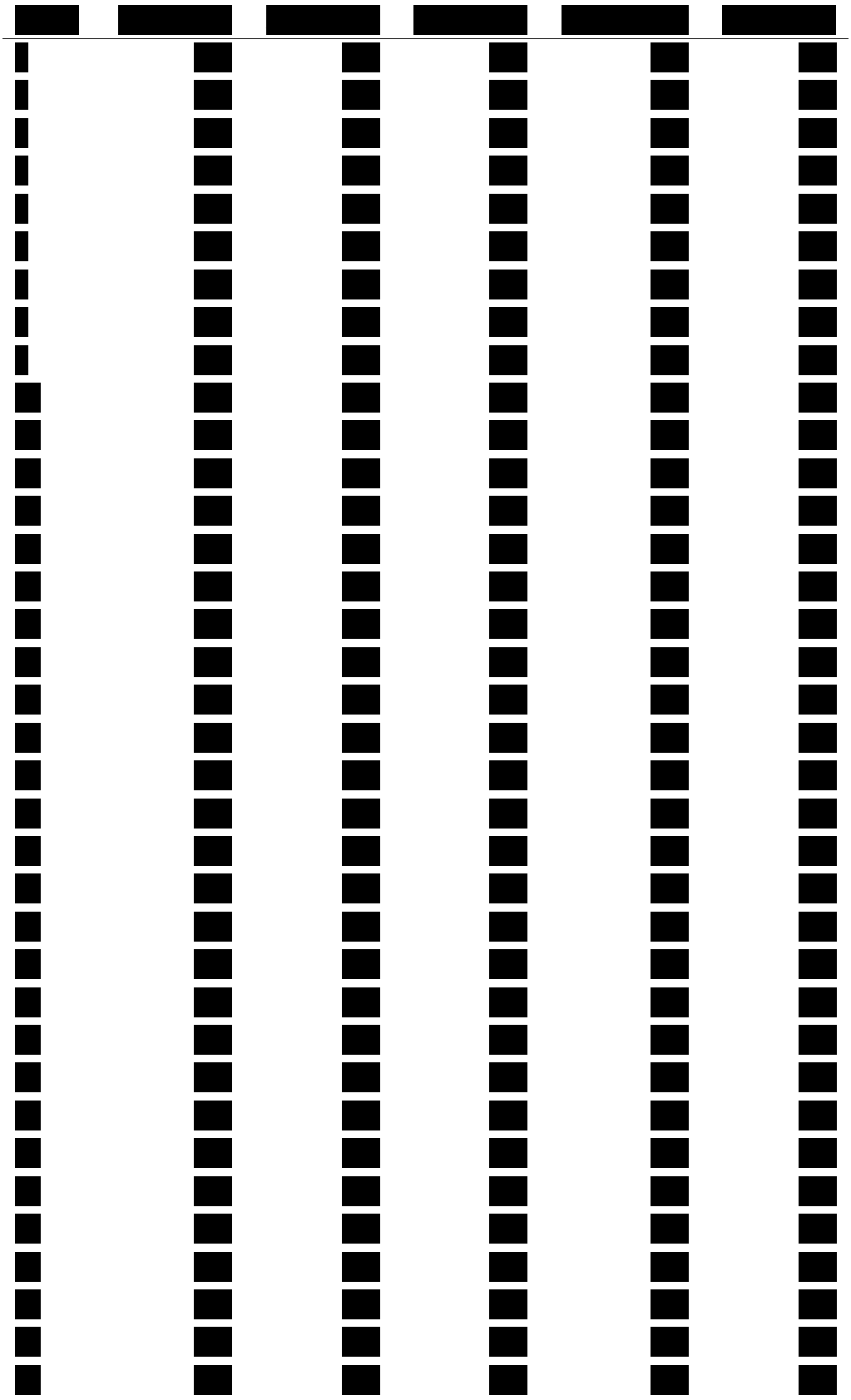
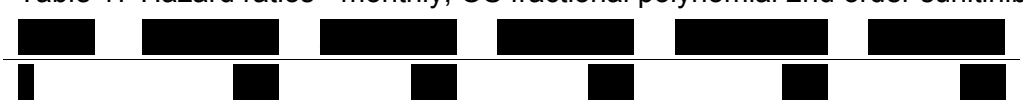


Table 17 Hazard ratios - monthly, OS fractional polynomial 2nd order sunitinib



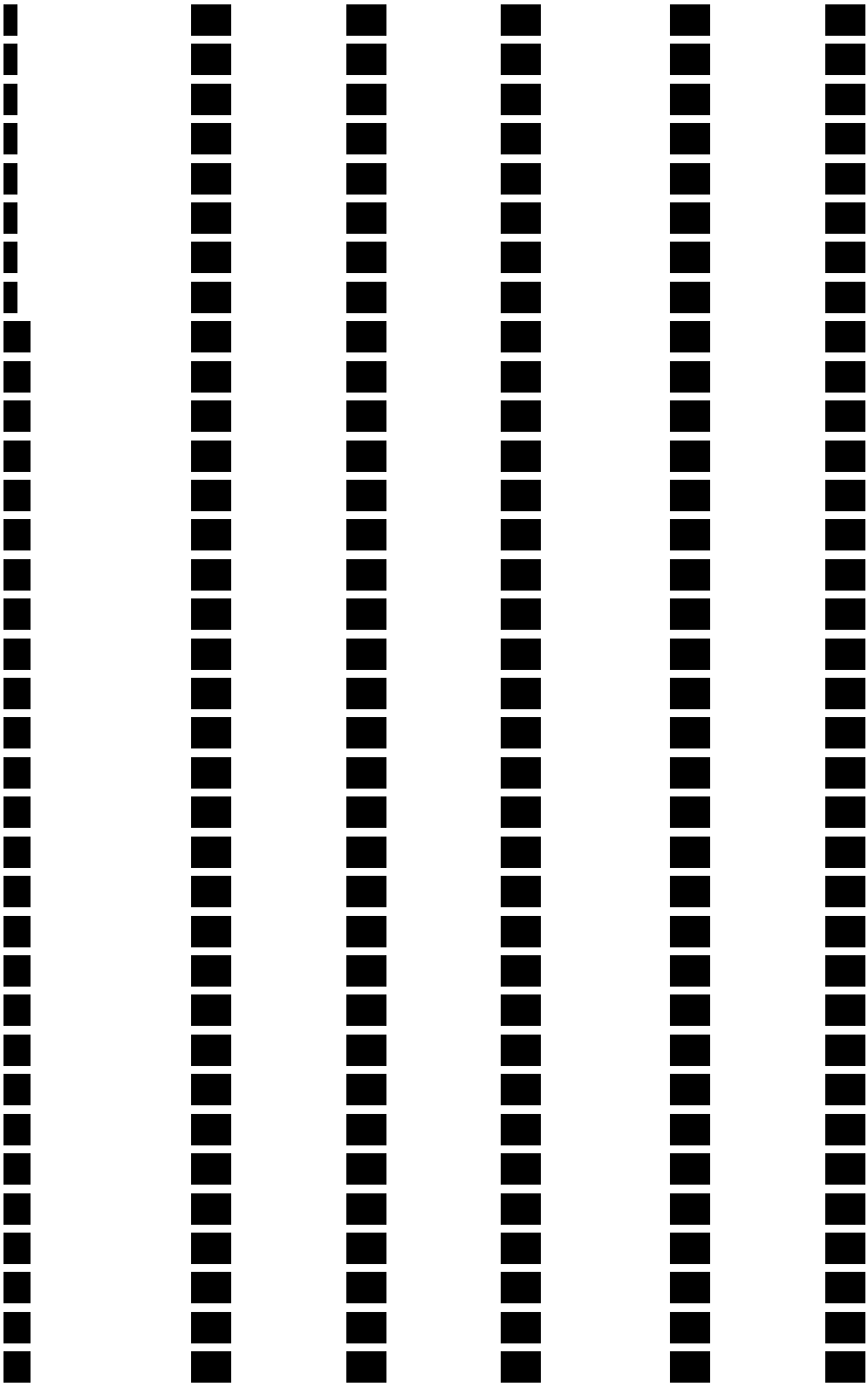


Table 18 Statistical fits – DIC (random effect fractional polynomial models)

FP	parameters	OS	PFS
1st	P=0	1475.2	1945.6
1st	P=1	1490.3	1943.6
1st	P=0.5	1485.0	1947.4

1st	P=-1	1450.3	1910.6
1st	P=-0.5	1461.8	1932.3
2nd	P1=-0.5, P2=0	1452.1	1854.1
2nd	P1=-1, P2=0	1450.3	1840.5
2nd	P1=-1, P2=-1	1449.2	1824.1
2nd	P1=-1, P2=0.5	1451.1	1850.4
2nd	P1=-1, P2=1	1451.9	1858.2

A24.

Priority question. Please explain why you chose to test fractional polynomials with a relatively narrow range of powers (P1 and P2 in the range -1 to +1). Given that none of the overall survival curves appears to reflect the shape of the CABOSUN Kaplan–Meier curves shown in Figure 6, did you consider trying other functional forms?

Response: Fractional polynomial (FP) models for hazard ratios (HR) are very flexible and fitting them imply the joint estimation of a large number of parameters for each model. In addition, no closed form [1] exists in this case to connect the hazard functions (expressed as fractional polynomials) with the survival probability, requiring an iterative numerical computation for survival (integral over the hazard, easily divergent). As a consequence, the joint estimation is very delicate for every (P1, P2) model and the lack of stability of the estimation algorithms typically causes very long run times, when numerical errors are not preventing this estimation all together. We therefore had to be strategic in our choices of which (P1, P2) to test. From our previous work on cabozantinib for treatment of second line aRCC, we knew that P=-1 and P1=P2=-1 provided the best fits out of the models that we tested. We therefore used these values as a starting point for the 1st-line indication. Another guiding principle we used in our choice of which (P1, P2) values to test is that smaller values of P1 and P2 should be preferred. The reason is that polynomials are famous for becoming “wigglier” and “wigglier” as their degree increases, and of course hazard functions are not.

A25.

Page 89 of the company submission states that “The FP 1st order models have lower DIC than the FP 2nd models, and therefore were not included in the scenario analyses”. We note from Table 24 that the fractional polynomials 1st order models have *higher* DIC values than the 2nd order models. Is this a typographical error?

Response: Correct; this is a typographical error. Page 8 should say “*The FP 1st order models have higher DIC than the FP 2nd models, and therefore were not included in the scenario analyses*”.

[1] In mathematics closed form is a mathematical expression that can be evaluated in a finite number of operations.

A26.

Priority question. We are interested to know whether the results of the ITC for the comparison of cabozantinib with sunitinib and pazopanib would differ if the wider primary evidence network shown in Figure 9 was used, rather than the restricted network shown in Figure 11. Please provide full results (random and fixed effect) of the ITC for the primary evidence network using the Ouwens et al and fractional polynomial methods, accompanied by an assessment of statistical and clinical heterogeneity, inconsistency, and methodological study quality.

A1. **Response**: When using the wider primary network of studies, there is a clinical heterogeneity present. TARGET study (sorafenib versus placebo) is in the wider network, and this study was conducted in a mostly pre-treated population.[3] Differences in extent of cross-over and/or subsequent therapies are also present. For example, in TIVO-1 study, more patients received therapy in sorafenib arm than in tivozanib arm, leading to potential under- of tivozanib OS.[4] All random and fixed effect models (from Figure 71 to Figure 182) suggest that cabozantinib significantly increases PFS compared with sunitinib and pazopanib, and that overall survival was longest for cabozantinib regardless of whether the model was run on wider or narrower network. The fixed effect Ouwens et al models are provided from Figure 71 to Figure 86. The random effect Ouwens et al. models are provided from Figure 87 to Figure 102. The fixed effect fractional polynomial models are provided from Figure 103 ***OS fractional polynomial model survival curve – fixed-effect model (P=0)***

[3] Negrier S, Jager E, Porta C, et al. Efficacy and safety of sorafenib in patients with advanced renal cell carcinoma with and without prior cytokine therapy, a subanalysis of TARGET. Med Oncol. 2010;27(3):899-906.
[4] Motzer RJ, Nosov D, Eisen T, et al. Tivozanib versus sorafenib as initial targeted therapy for patients with metastatic renal cell carcinoma: results from a phase III trial. J Clin Oncol. 2013;31(30):3791-3799.

Figure 104 OS fractional polynomial model HR plot – fixed-effect model (P=0)

Figure 105 OS fractional polynomial model survival curve – fixed-effect model (P=1)

Figure 106 OS fractional polynomial model HR plot – fixed-effect model (P=1)

Figure 107 OS fractional polynomial model survival curve – fixed-effect model (P=0.5)

Figure 108 OS fractional polynomial model HR plot – fixed-effect model (P=0.5)

Figure 109 OS fractional polynomial model survival curve – fixed-effect model (P=-1)

Figure 110 OS fractional polynomial model HR plot – fixed-effect model (P=-1)

Figure 111 OS fractional polynomial model survival curve – fixed-effect model (P=-0.5)

Figure 112 OS fractional polynomial model HR plot – fixed-effect model (P=-0.5)

Table 51 OS fractional polynomial model monthly HRs – random-effect model BEV+IFN

[redacted]

Table 52 OS fractional polynomial model monthly HRs – random-effect model IFN

[redacted]

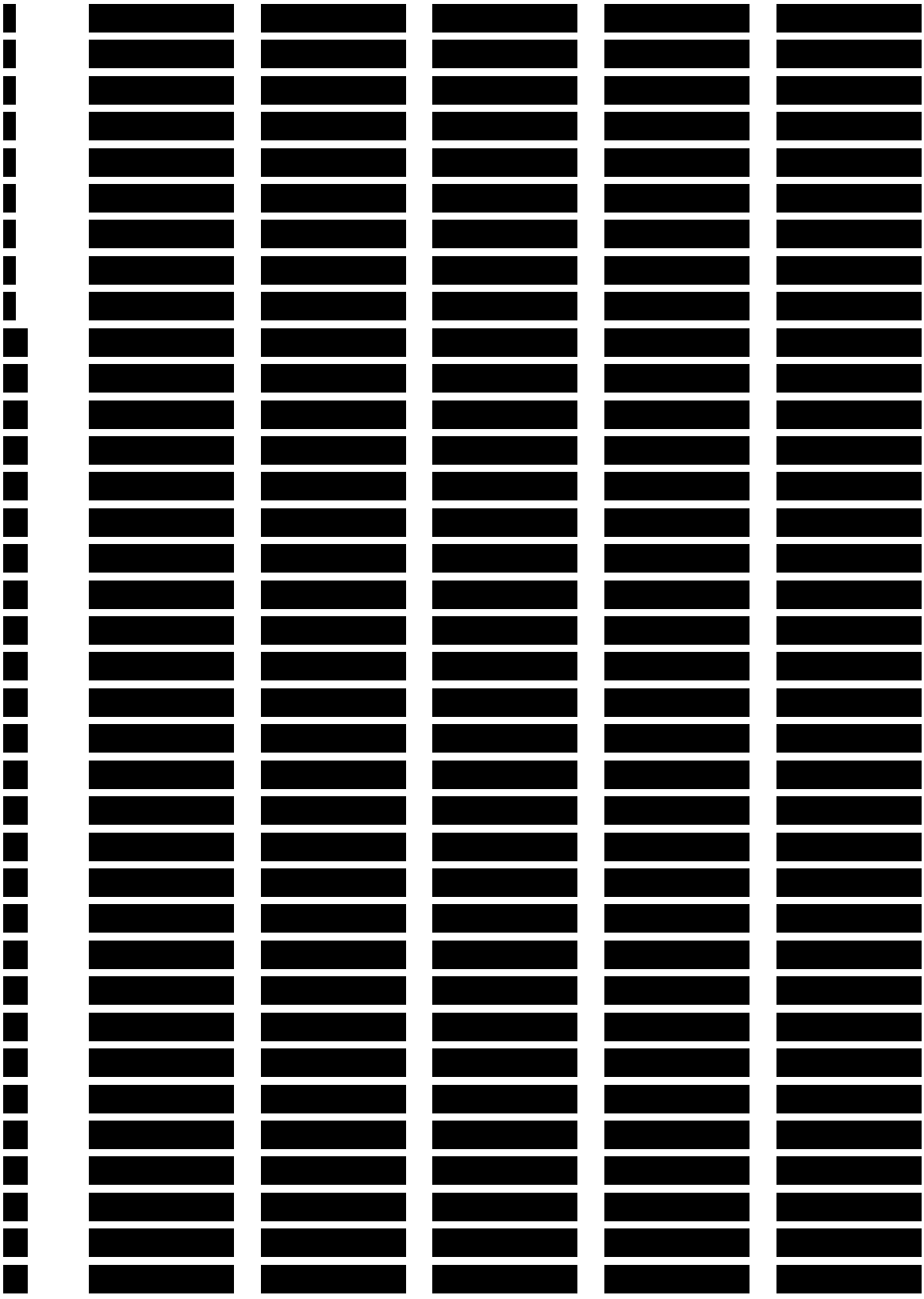
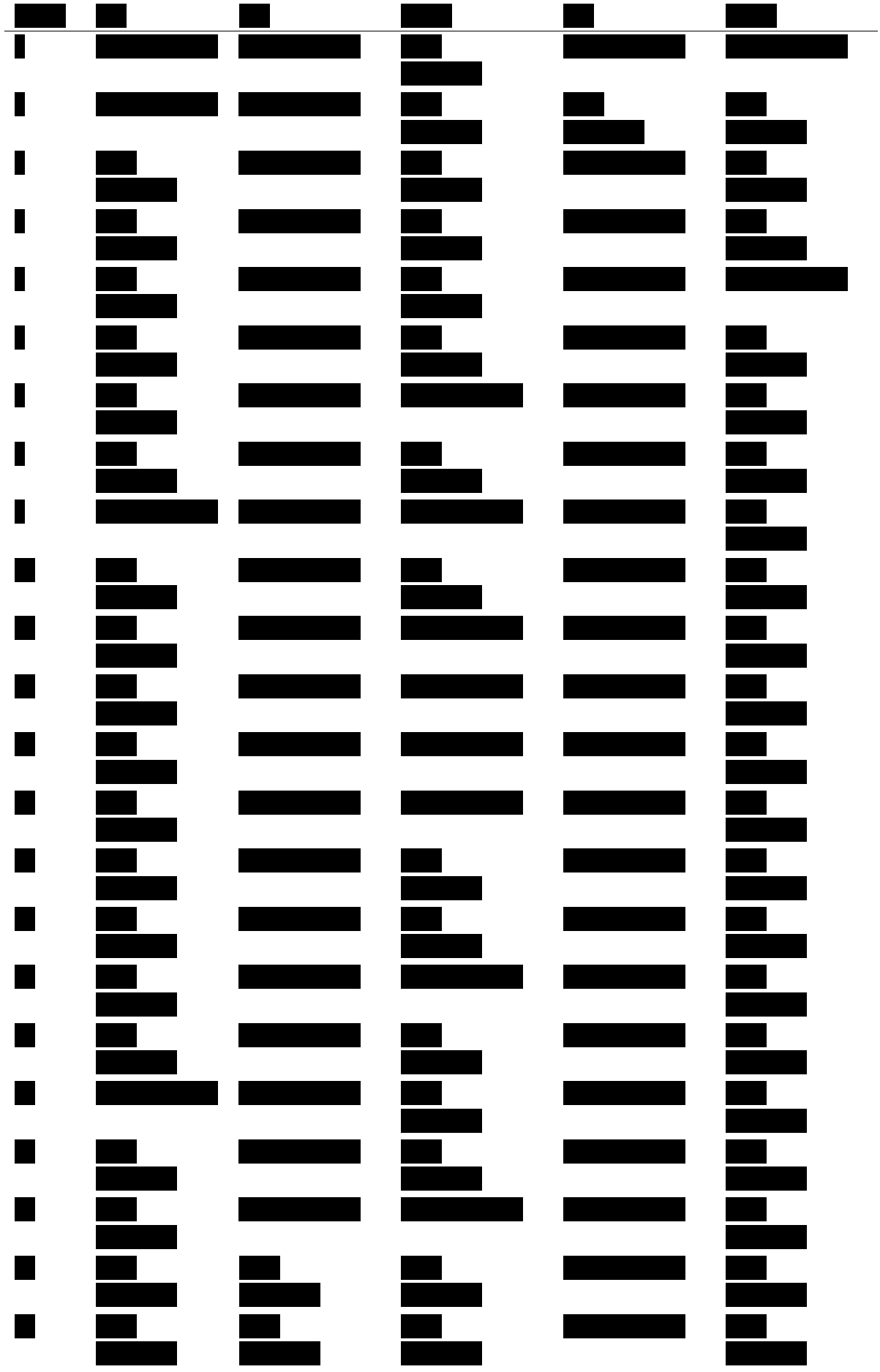


Table 53 OS fractional polynomial model monthly HRs – random-effect model pazopanib

[redacted]

Table 54 OS fractional polynomial model monthly HRs – random-effect model sorafenib



[redacted]

Table 55 OS fractional polynomial model monthly HRs – random-effect model sunitinib

[redacted]

[redacted]

Table 56 OS fractional polynomial model monthly HRs – random-effect model temsirolimus

[redacted]

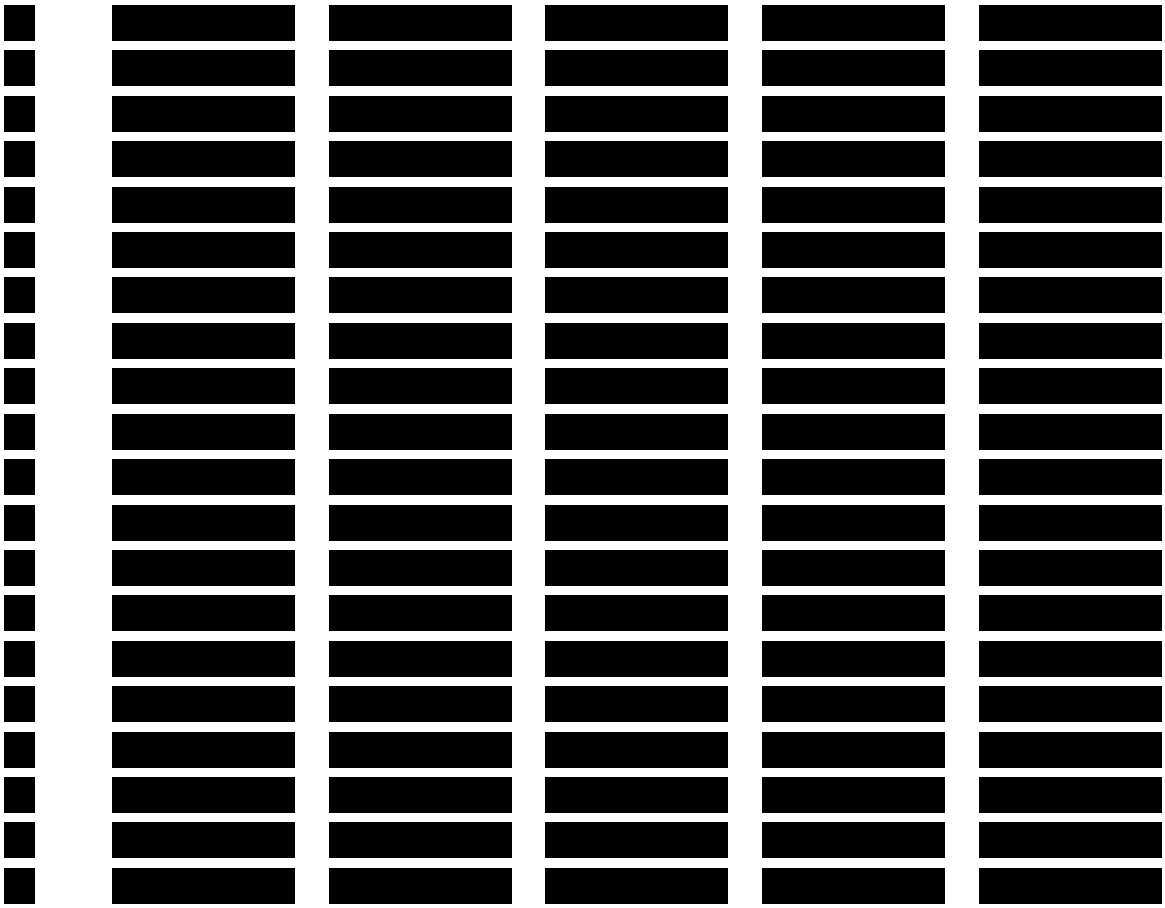
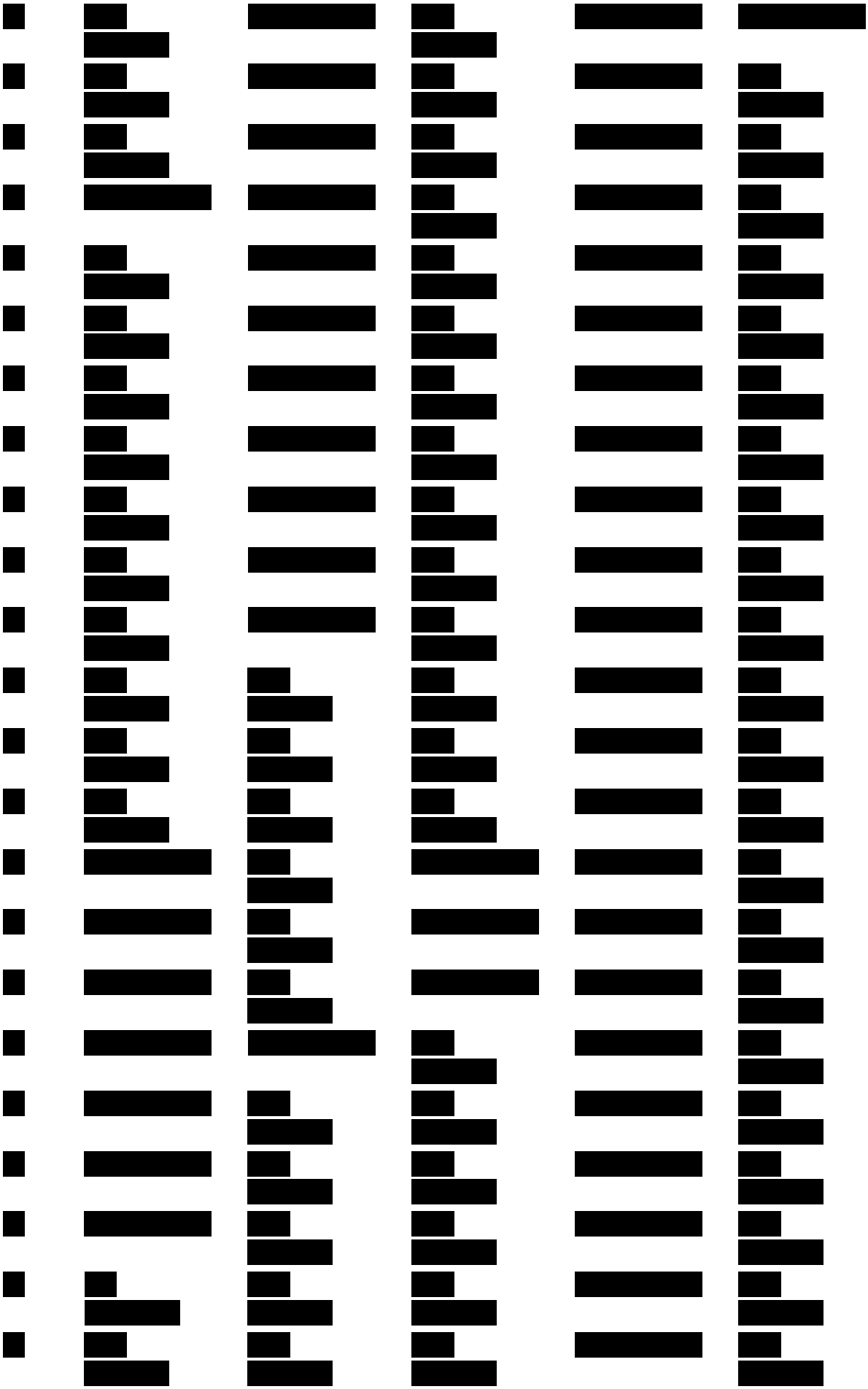
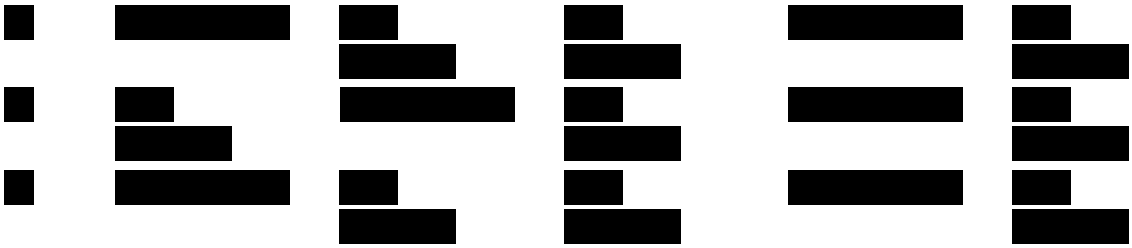


Table 57 OS fractional polynomial model monthly HRs – random-effect model tivozanib





and Figure 142, and random effect fractional polynomial models from Figure 143 to Figure 182. For credible intervals, please see response to question A23.

Fixed effect Ouwens et al. models

OS fixed effect Ouwens et al.

Figure 71 OS Ouwens model survival curve – fixed-effect model (log-logistic)

Figure 72 OS Ouwens model HR plot – fixed-effect model (log-logistic)

Figure 73 OS Ouwens model survival curve – fixed-effect model (exponential)

Figure 74 OS Ouwens model HR plot – fixed-effect model (exponential)

Figure 75 OS Ouwens model survival curve – fixed-effect model (Weibull)

Figure 76 OS Ouwens model HR plot – fixed-effect model (Weibull)

Figure 77 OS Ouwens model survival curve – fixed-effect model (log-normal)

Figure 78 OS Ouwens model HR plot – fixed-effect model (log-normal)

Table 19 OS Ouwens model monthly HRs – fixed-effect model BEV+IFN

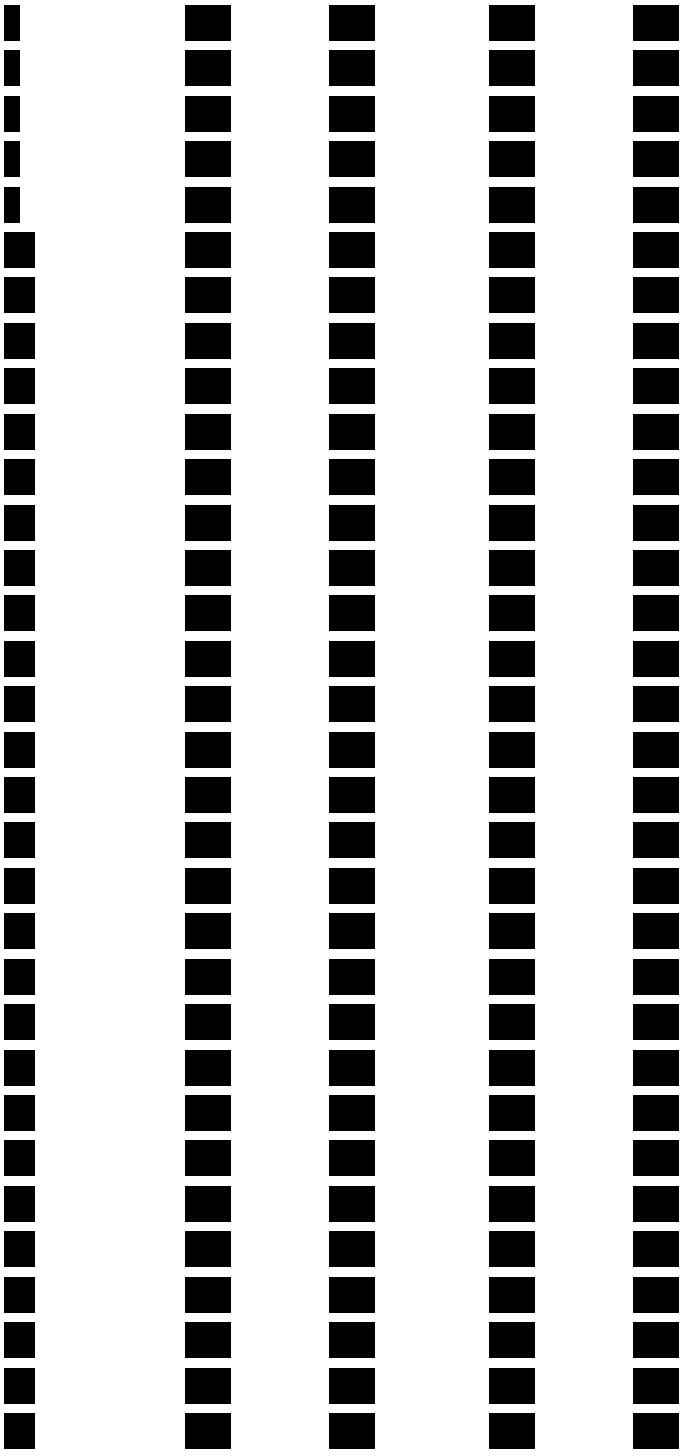
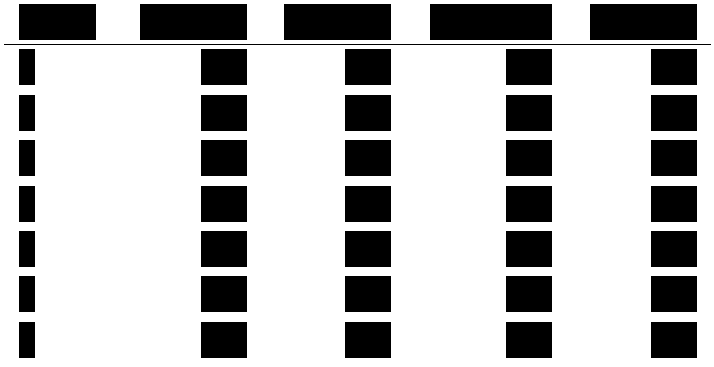


Table 20 OS Ouwens model monthly HRs – fixed-effect model IFN



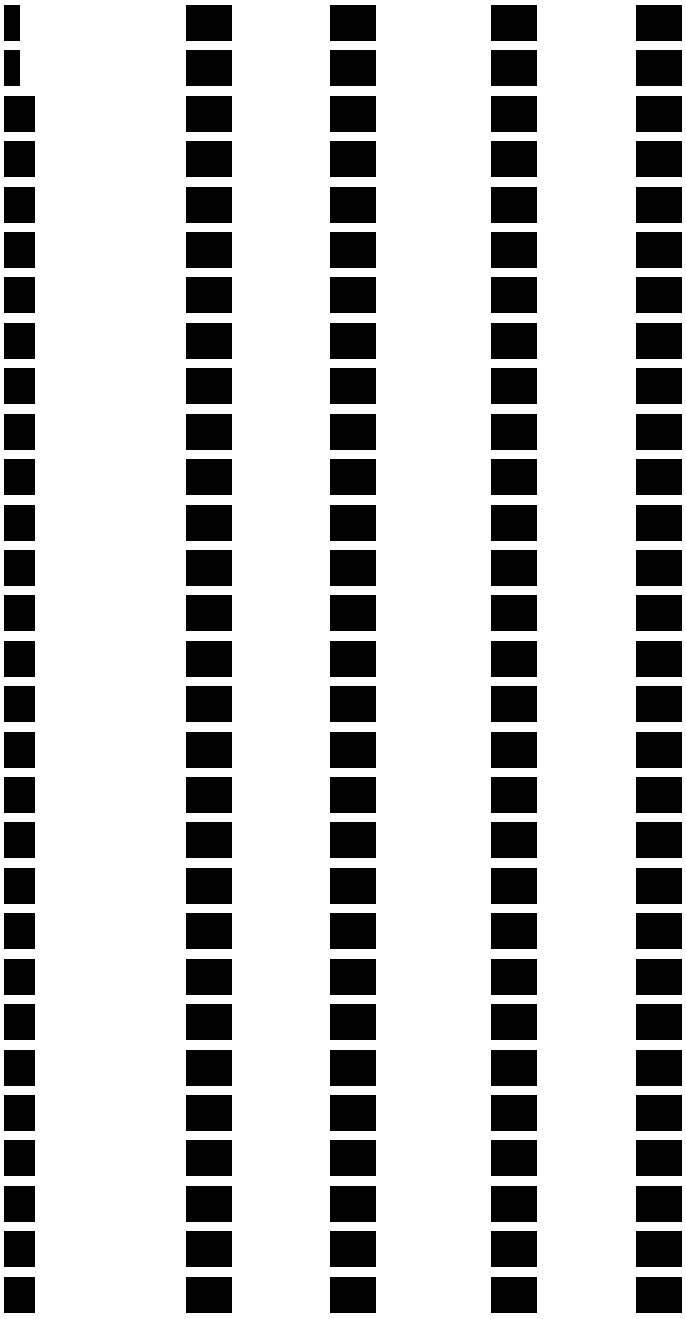
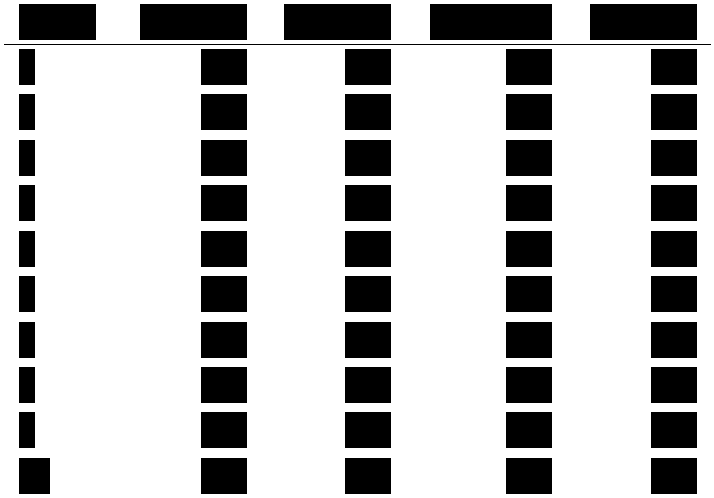


Table 21 OS Ouwens model monthly HRs – fixed-effect model pazopanib



[redacted]

Table 22 OS Ouwens model monthly HRs – fixed-effect model sorafenib

[redacted]

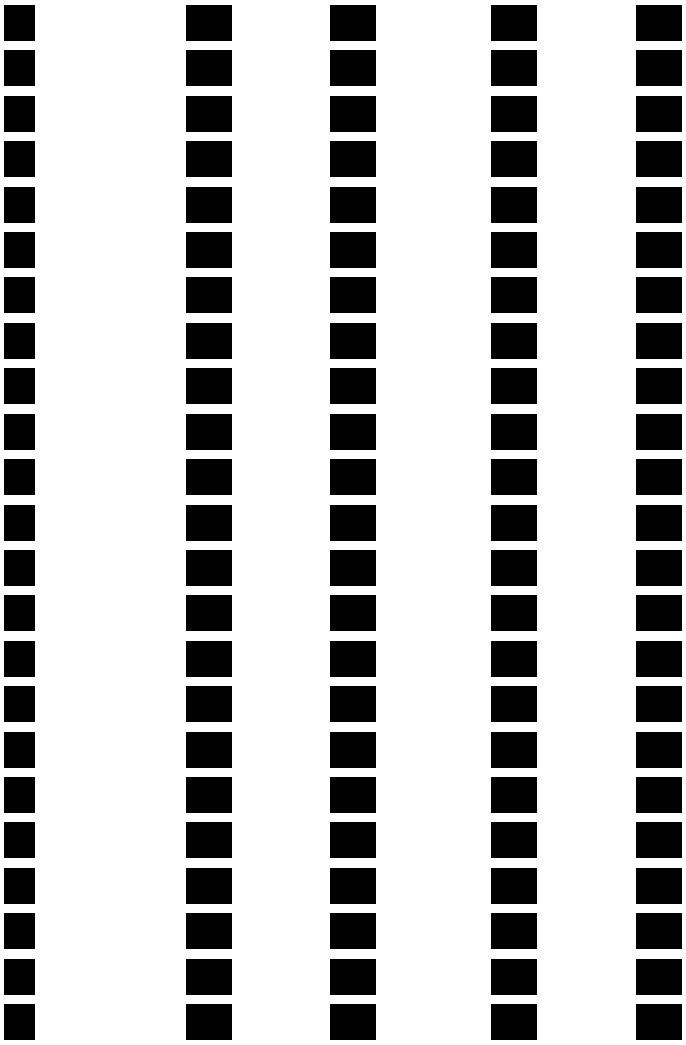
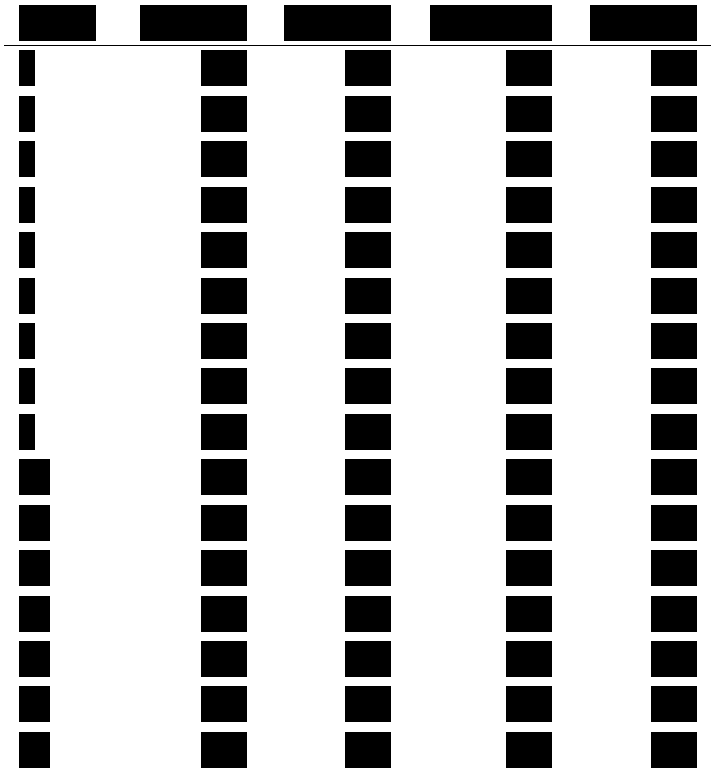


Table 23 OS Ouwens model monthly HRs – fixed-effect model sunitinib



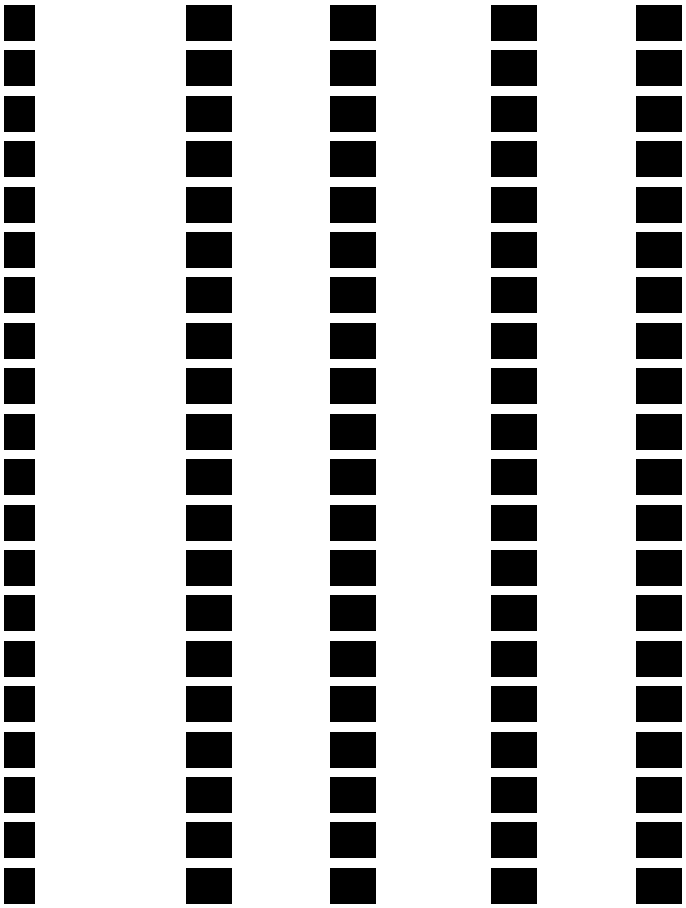
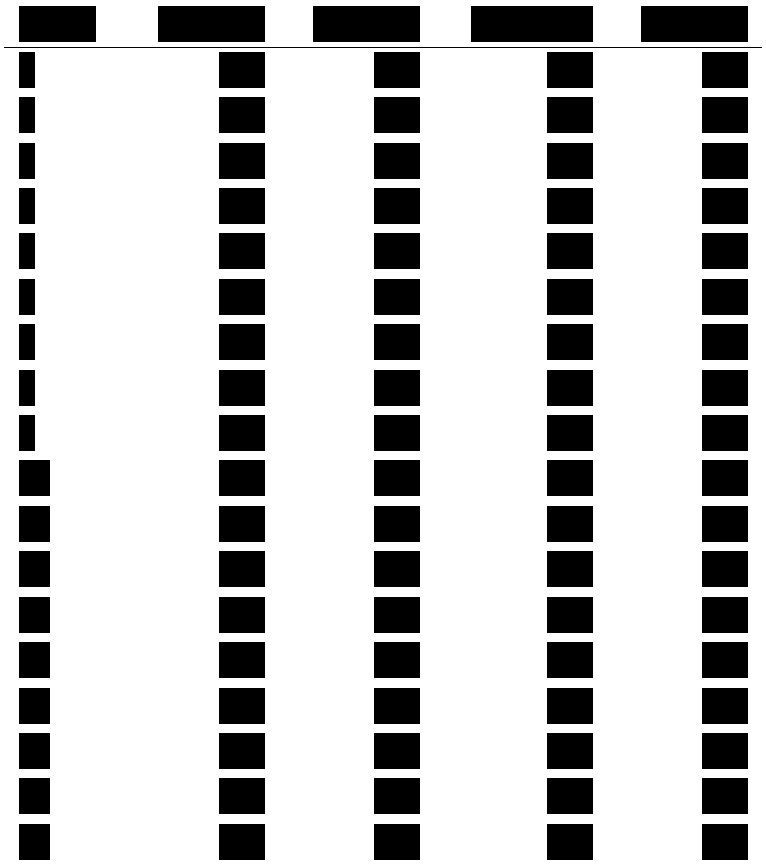


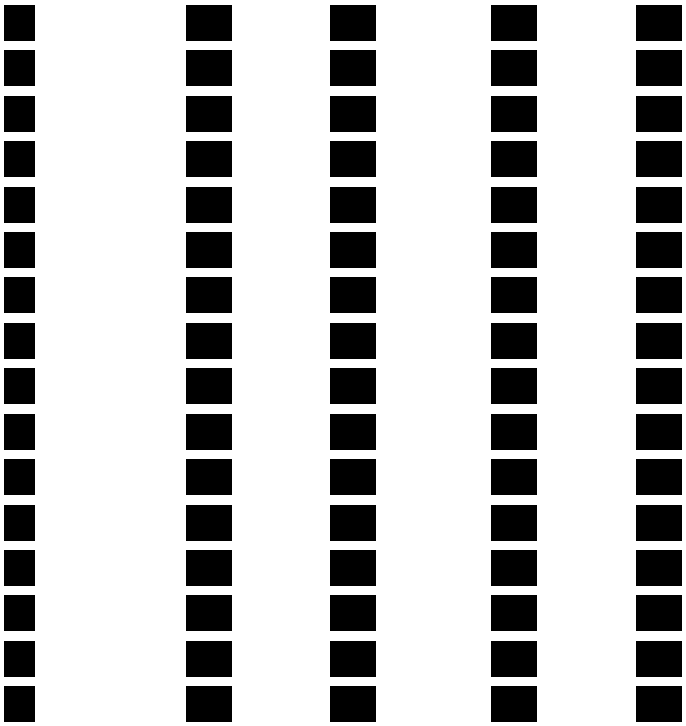
Table 24 OS Ouwens model monthly HRs – fixed-effect model temsirolimus



[redacted]

Table 25 OS Ouwens model monthly HRs – fixed-effect model tivozanib

[redacted]



PFS fixed effect Ouwens et al.

Figure 79 PFS Ouwens model survival curve – fixed-effect model (log-logistic)

Figure 80 PFS Ouwens model HR plot – fixed-effect model (log-logistic)

Figure 81 PFS Ouwens model survival curve – fixed-effect model (exponential)

Figure 82 PFS Ouwens model HR plot – fixed-effect model (exponential)

Figure 83 PFS Ouwens model survival curve – fixed-effect model (Weibull)

Figure 84 PFS Ouwens model HR plot – fixed-effect model (Weibull)

Figure 85 PFS Ouwens model survival curve – fixed-effect model (log-normal)

Figure 86 PFS Ouwens model HR plot – fixed-effect model (log-normal)

Table 26 PFS Ouwens model monthly HRs – fixed-effect model BEV+IFN

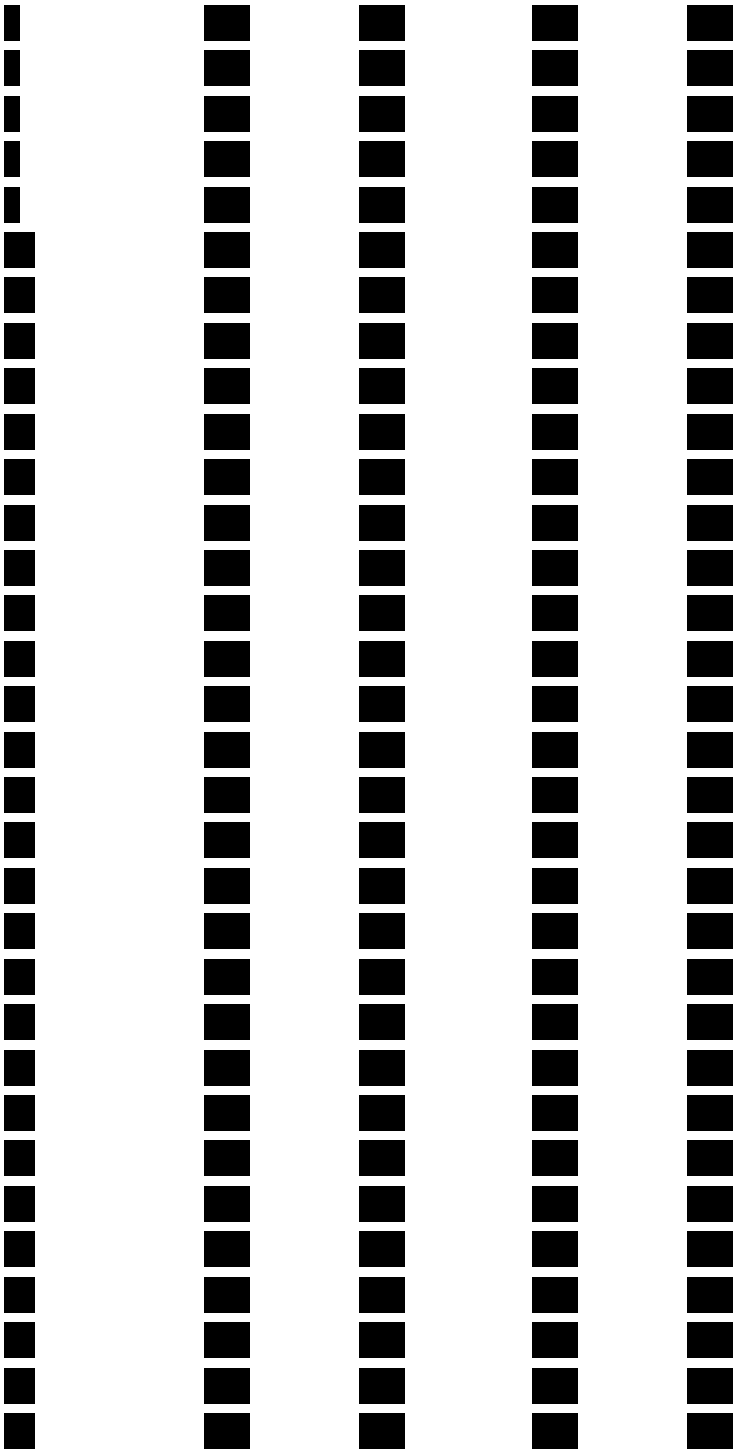
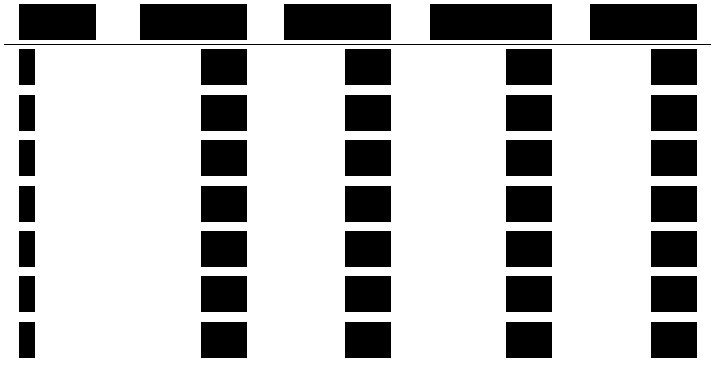


Table 27 PFS Ouwens model monthly HRs – fixed-effect model IFN



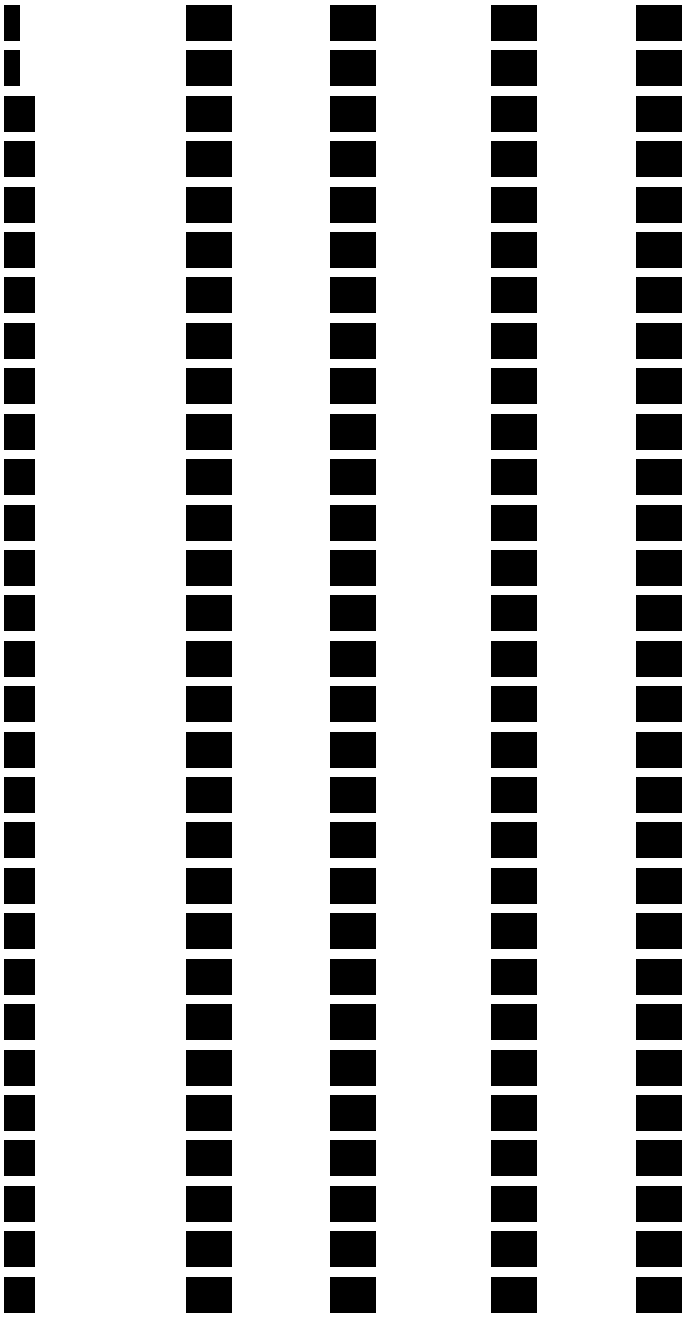
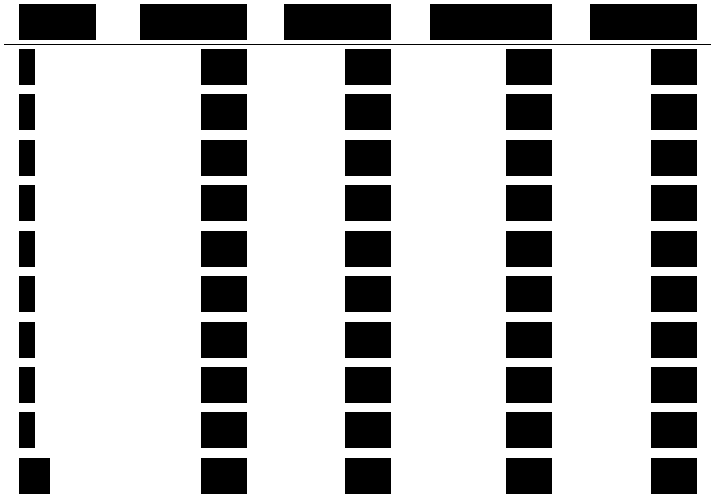


Table 28 PFS Ouwens model monthly HRs – fixed-effect model pazopanib



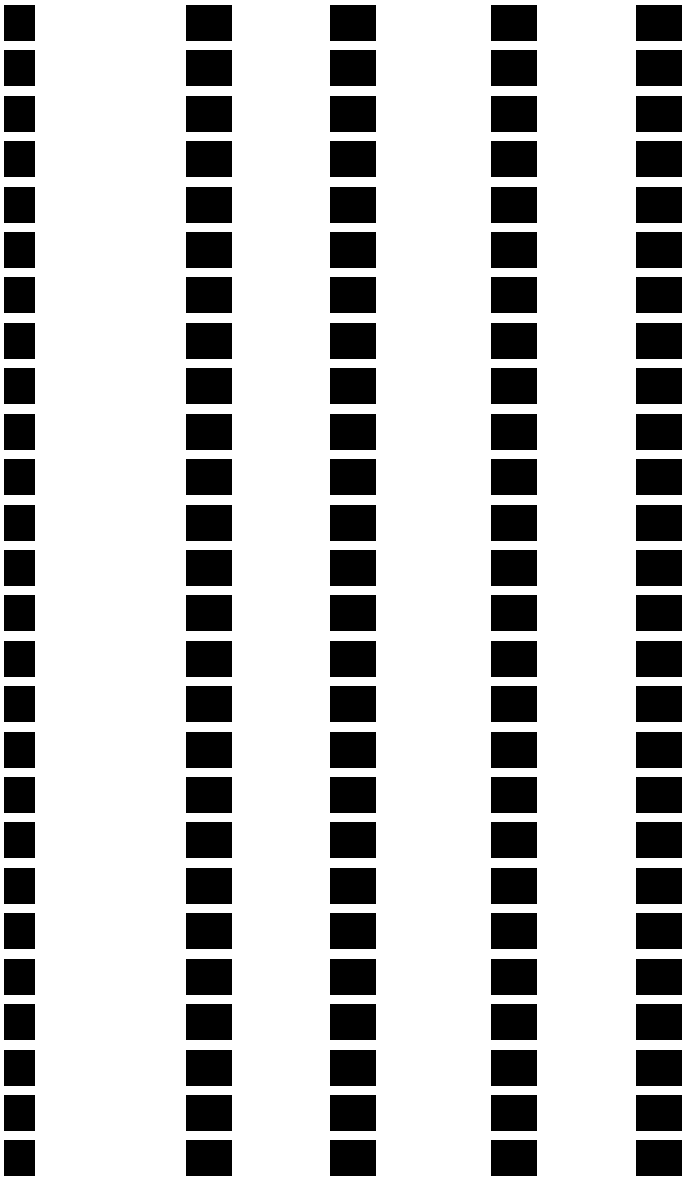
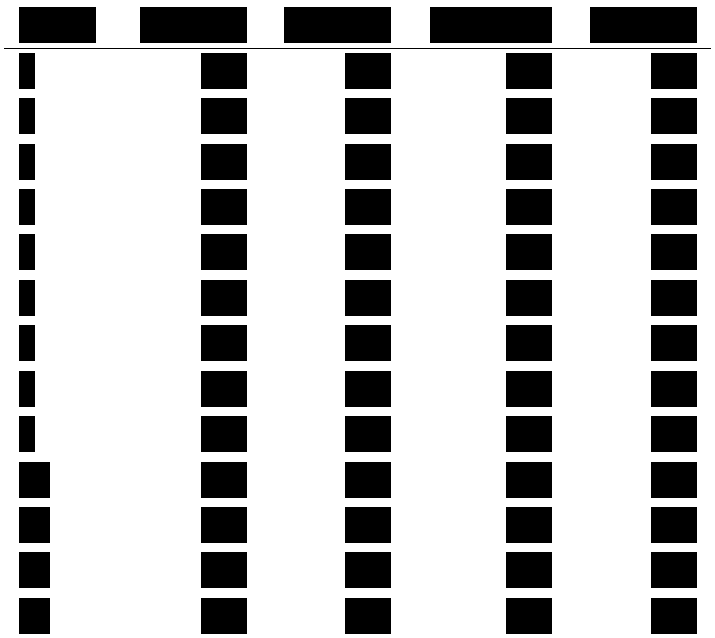


Table 29 PFS Ouwens model monthly HRs – fixed-effect model placebo



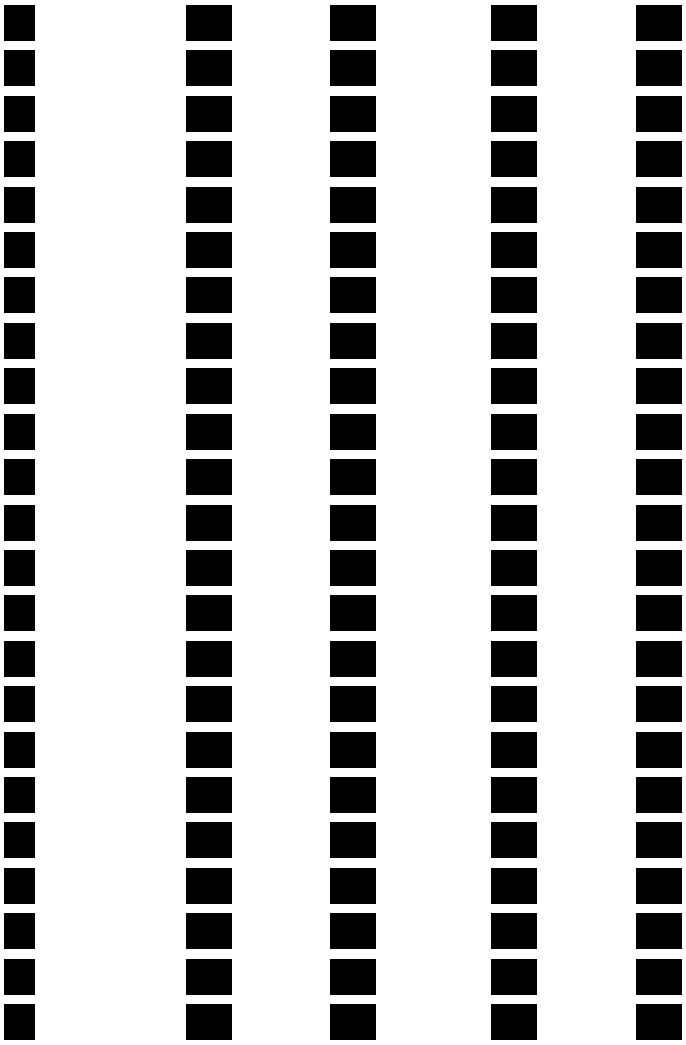
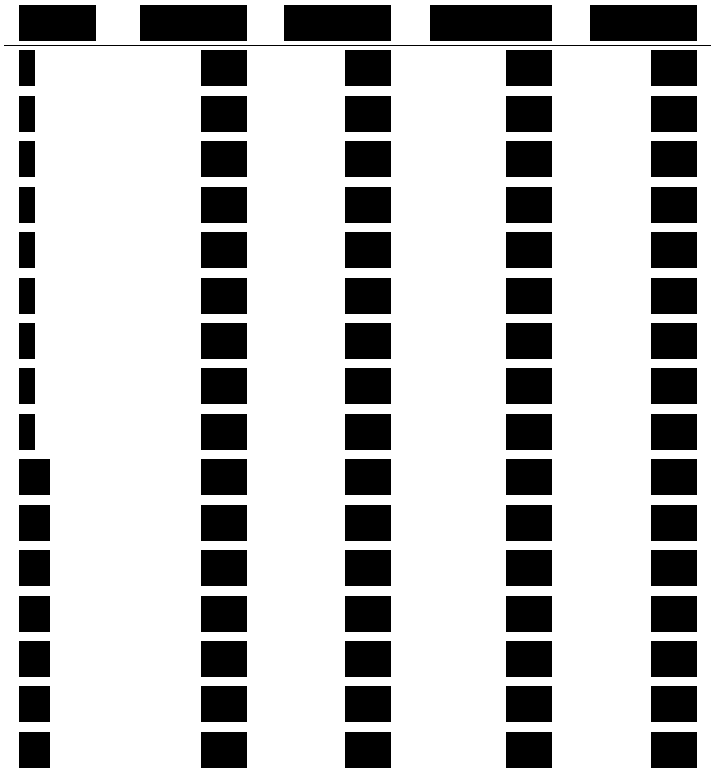


Table 30 PFS Ouwens model monthly HRs – fixed-effect model sorafenib



[redacted]

Table 31 PFS Ouwens model monthly HRs – fixed-effect model sunitinib

[redacted]

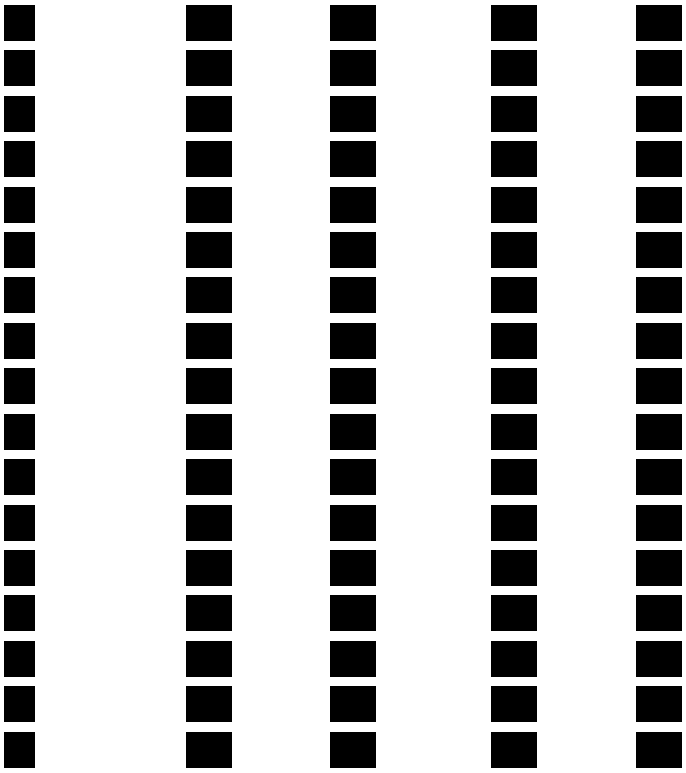
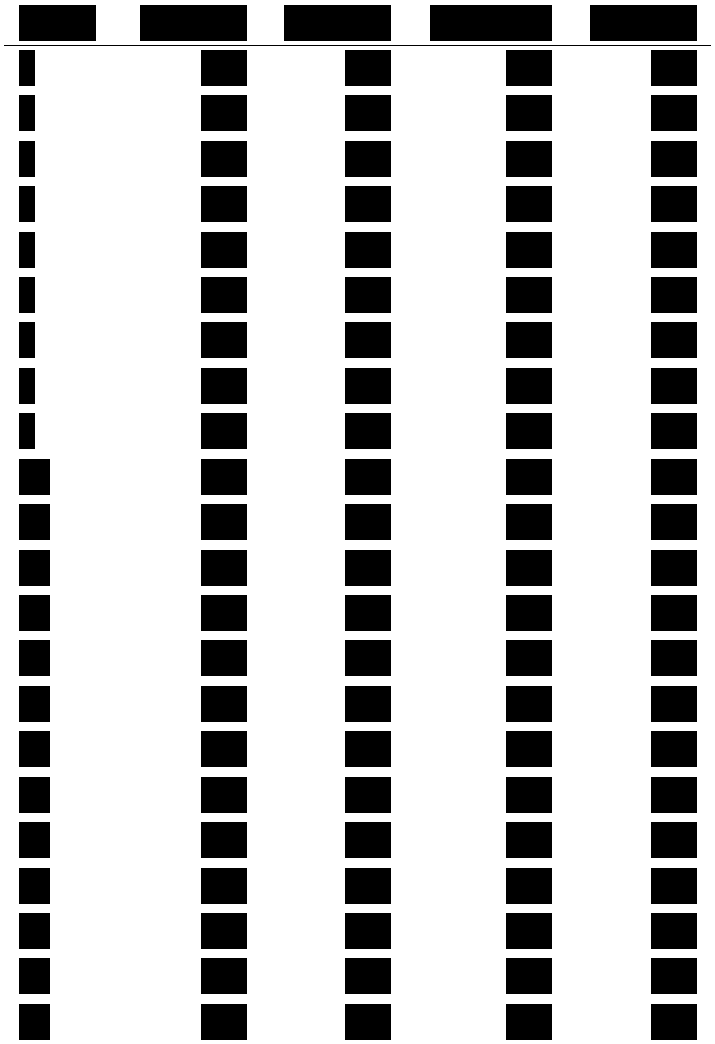


Table 32 PFS Ouwens model monthly HRs – fixed-effect model temsirolimus



[redacted]

Table 33 PFS Ouwens model monthly HRs – fixed-effect model tivozanib

[redacted]

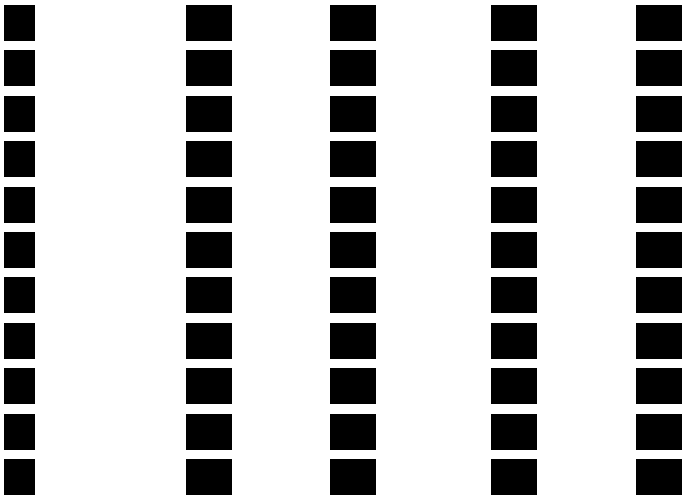


Table 34 Statistical fits – DIC Ouwens fixed-effect

parameters	OS	PFS
log-logistic	5662.1	9207.4
exponential	5776.7	9646.0
Weibull	5751.9	9567.8
lognormal	5611.4	9084.6

Random effect Ouwens et al. models

OS random effect Ouwens et al.

Figure 87 OS Ouwens model survival curve – random-effect model (log-logistic)

Figure 88 OS Ouwens model HR plot – random-effect model (log-logistic)

Figure 89 OS Ouwens model survival curve – random-effect model (exponential)

Figure 90 OS Ouwens model HR plot – random-effect model (exponential)

Figure 91 OS Ouwens model survival curve – random-effect model (Weibull)

Figure 92 OS Ouwens model HR plot – random-effect model (Weibull)

Figure 93 OS Ouwens model survival curve – random-effect model (log-normal)

Figure 94 OS Ouwens model HR plot – random-effect model (log-normal)

Table 35 OS Ouwens model monthly HRs – random-effect model BEV+IFN

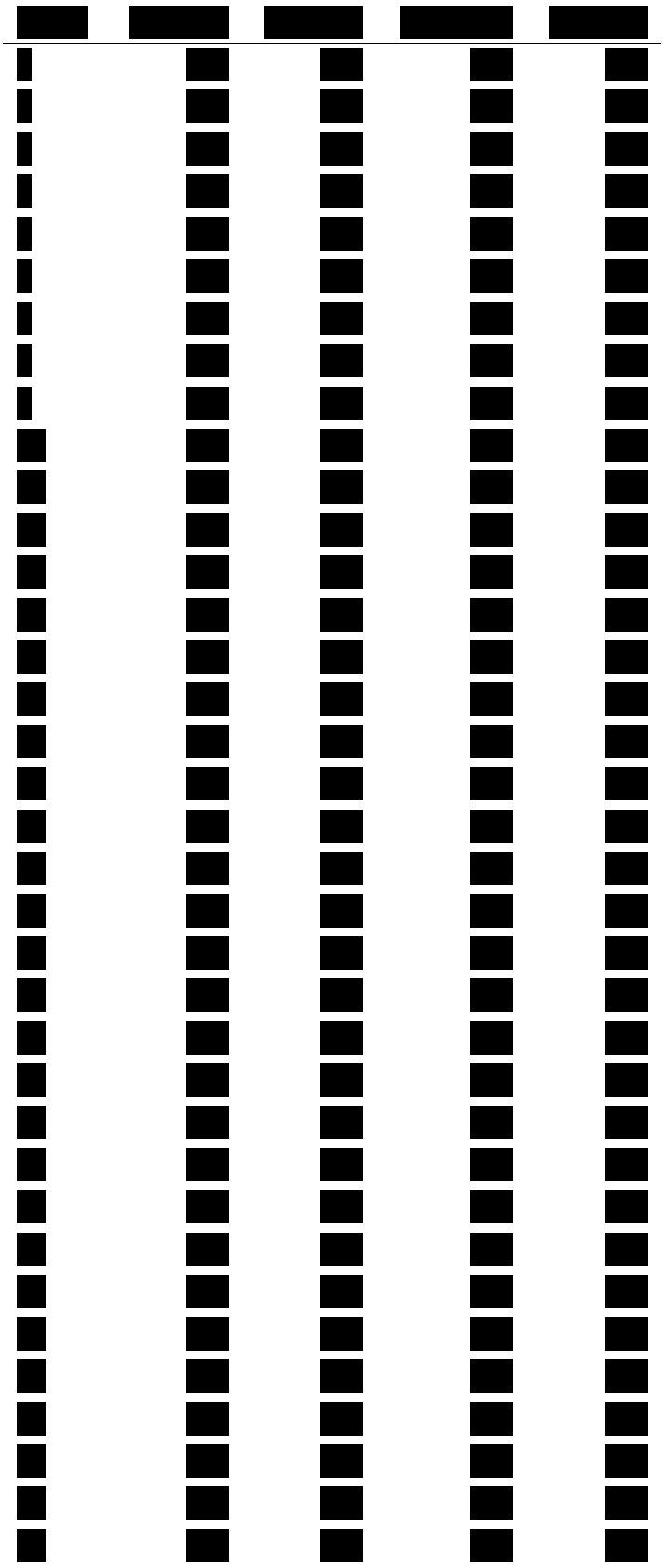


Table 36 OS Ouwens model monthly HRs – random-effect model IFN

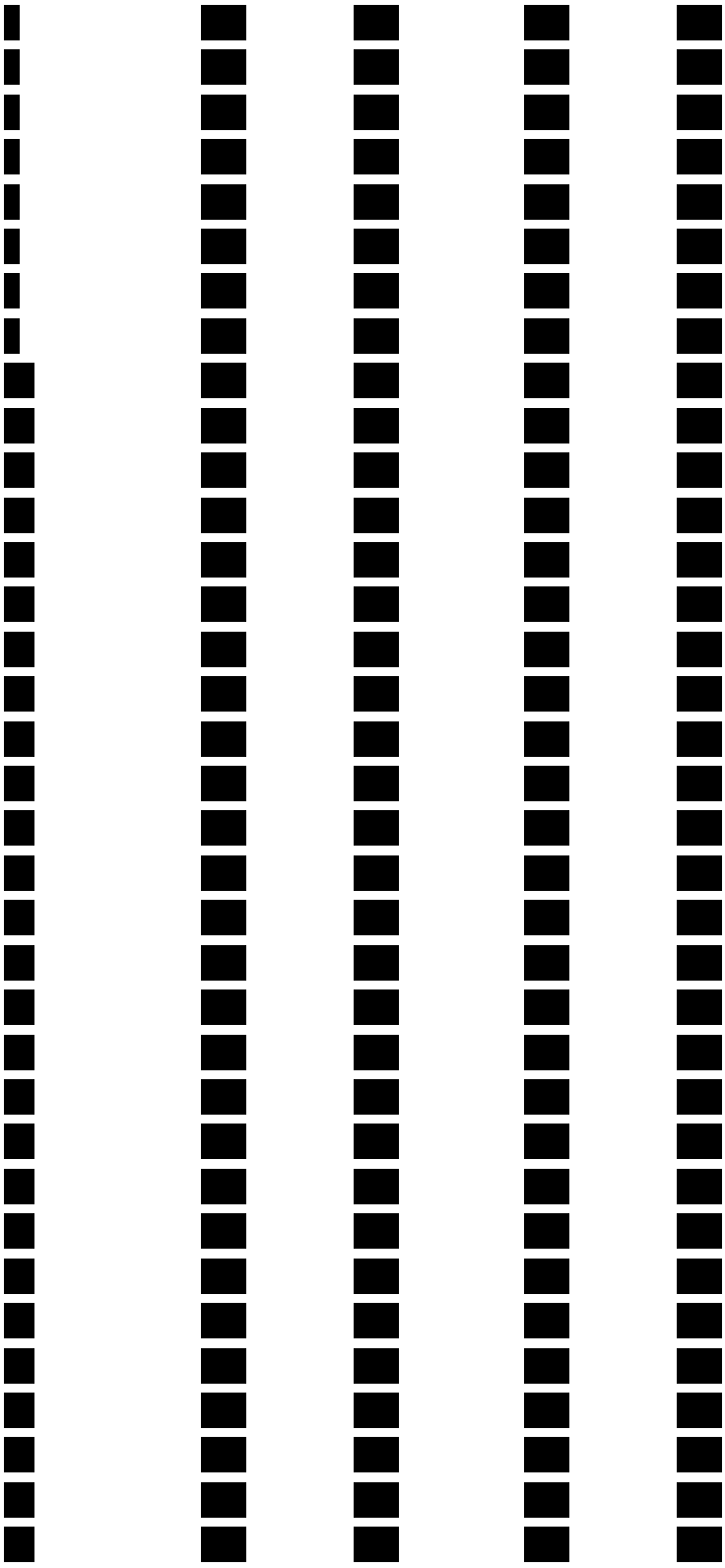
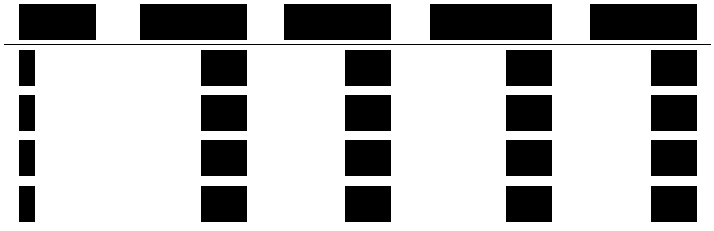


Table 37 OS Ouwens model monthly HRs – random-effect model pazopanib



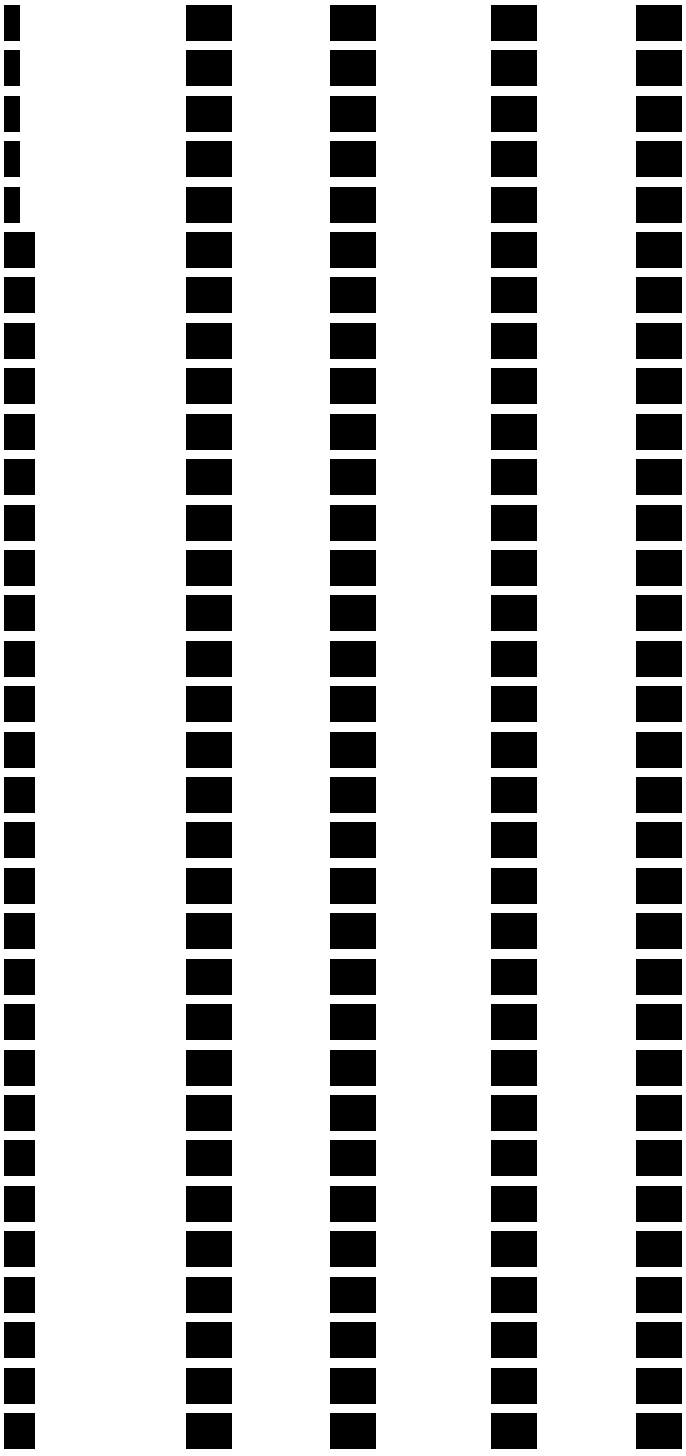
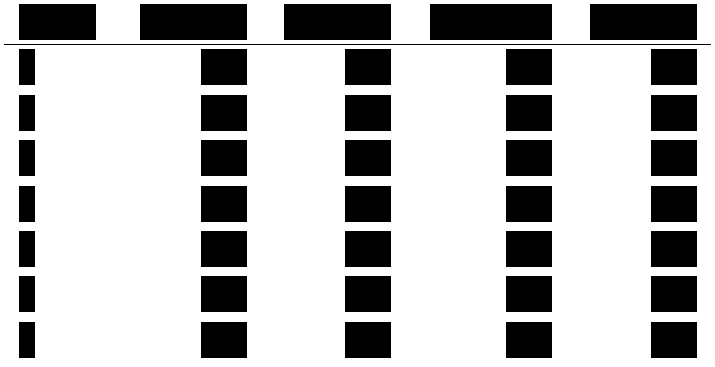


Table 38 OS Ouwens model monthly HRs – random-effect model sorafenib



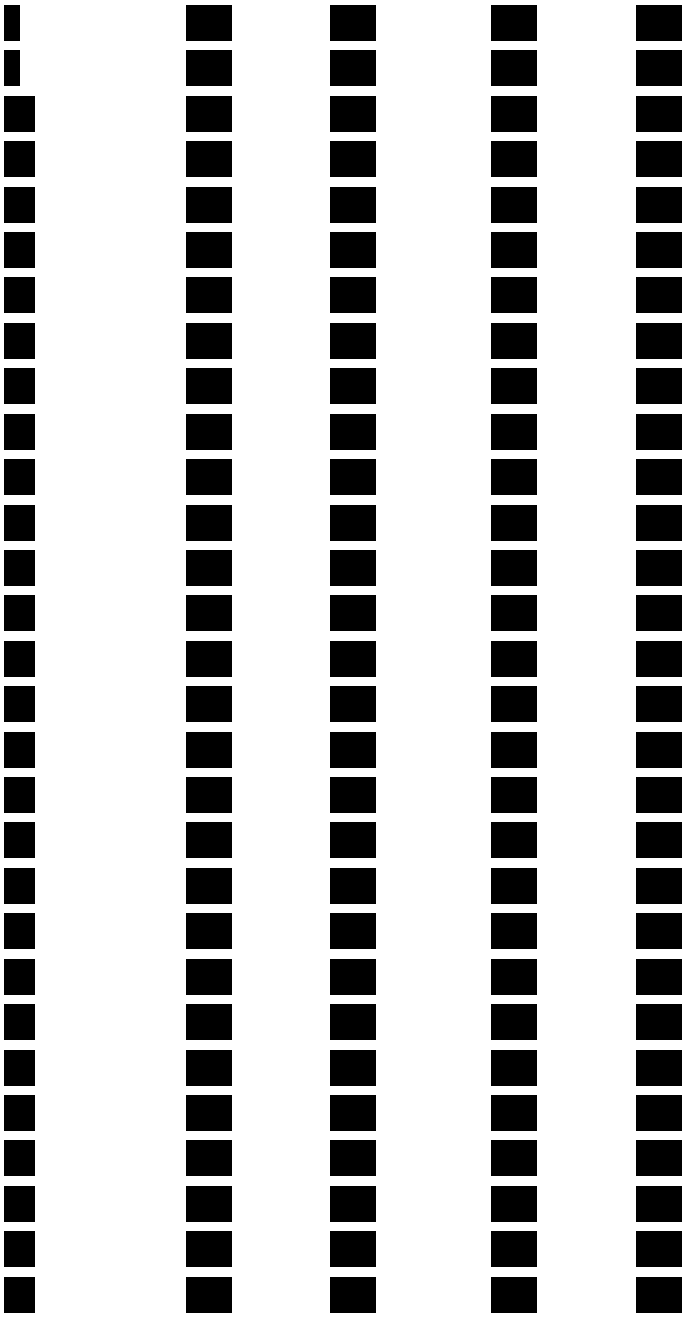
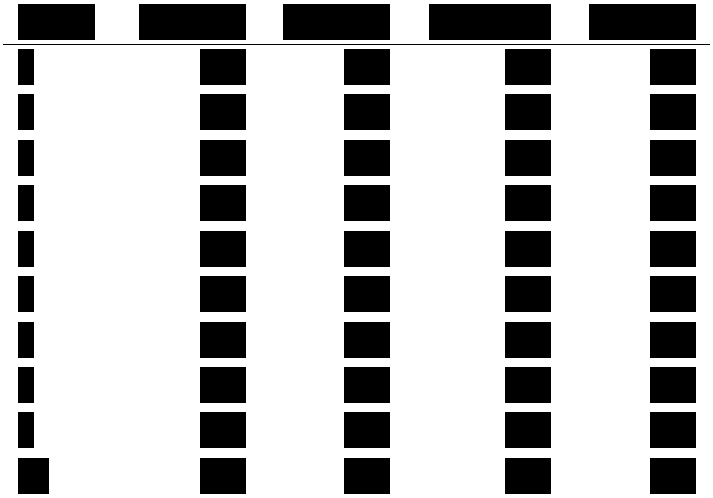


Table 39 OS Ouwens model monthly HRs – random-effect model sunitinib



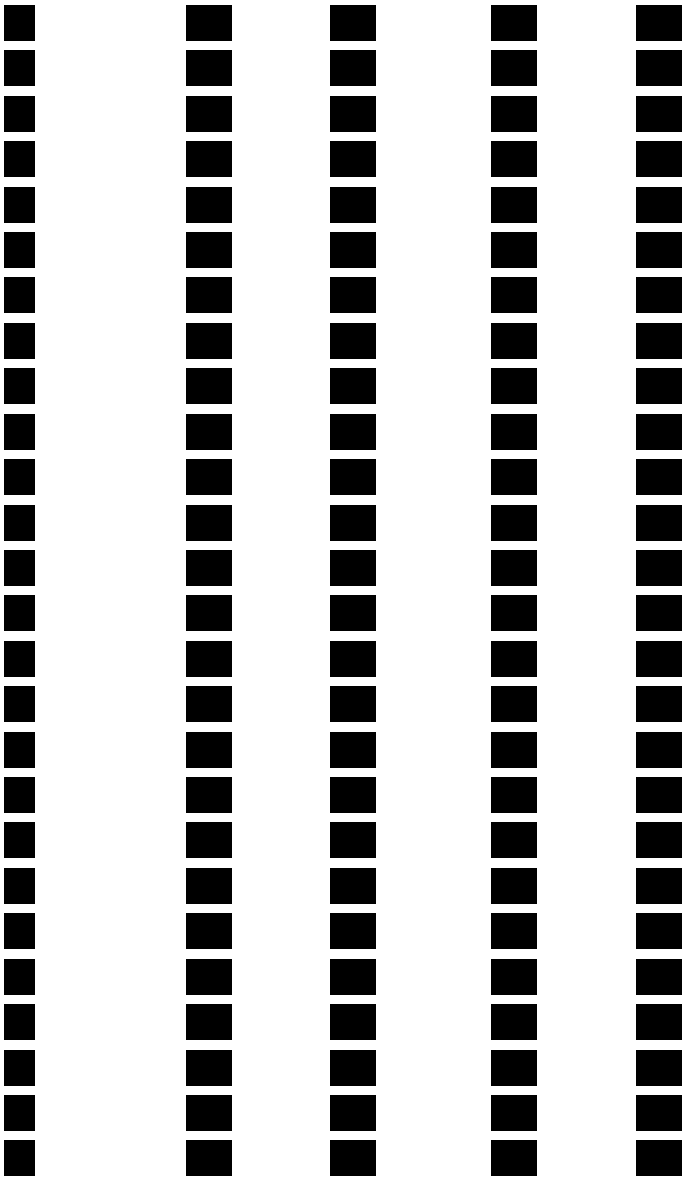
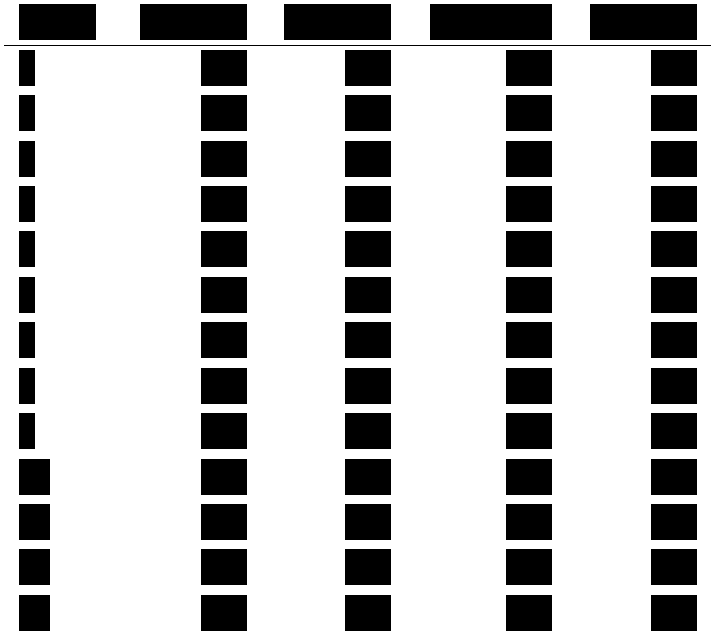


Table 40 OS Ouwens model monthly HRs – random-effect model temsirolimus



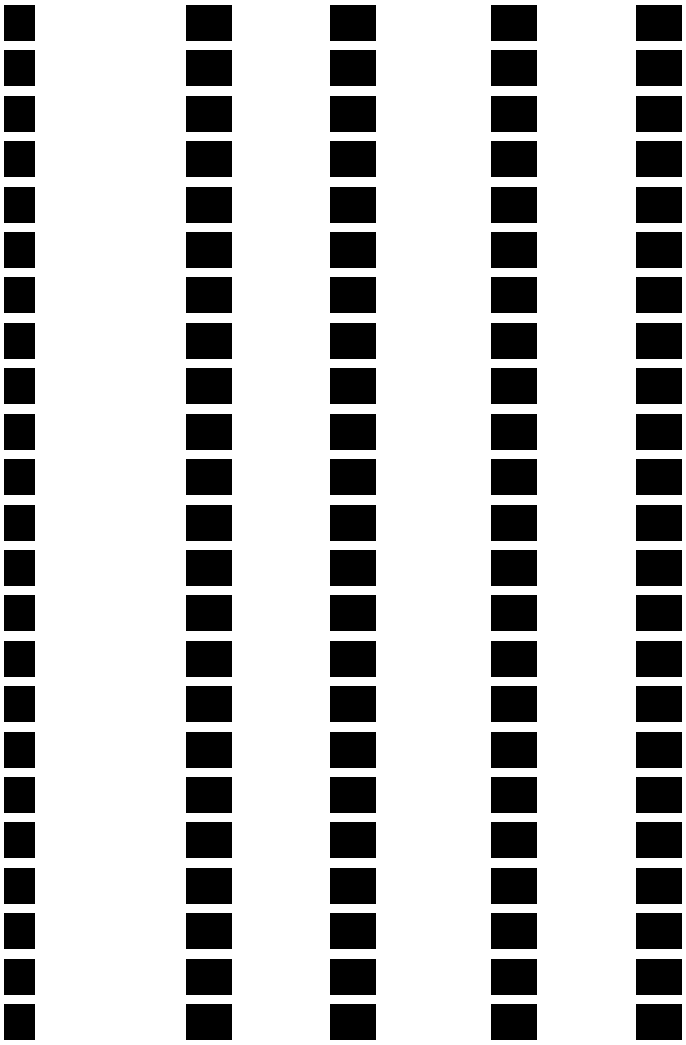
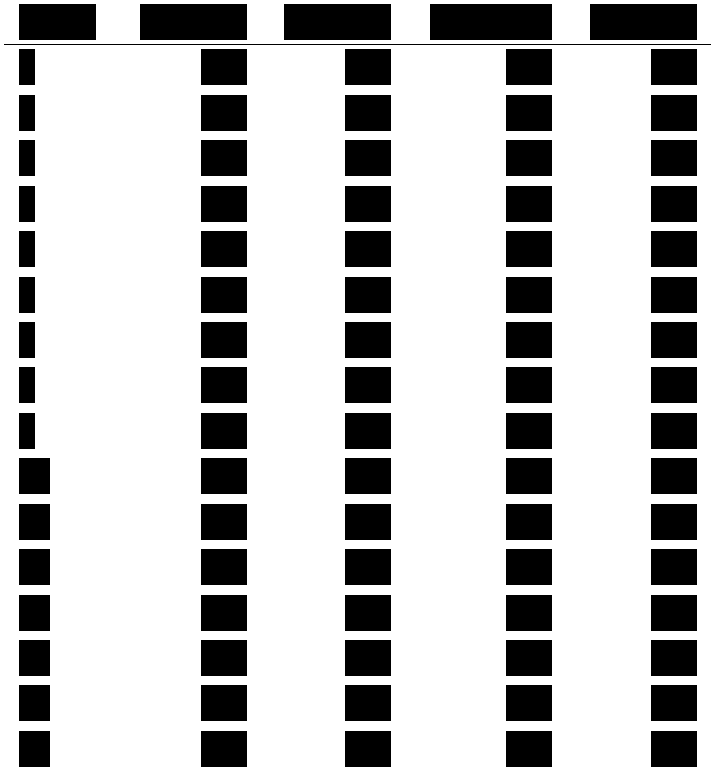


Table 41 OS Ouwens model monthly HRs – random-effect model tivozanib



PFS random effect Ouwens et al.

Figure 95 PFS Ouwens model survival curve – random-effect model (log-logistic)

Figure 96 PFS Ouwens model HR plot – random-effect model (log-logistic)

Figure 97 PFS Ouwens model survival curve – random-effect model (exponential)

Figure 98 PFS Ouwens model HR plot – random-effect model (exponential)

Figure 99 PFS Ouwens model survival curve – random-effect model (Weibull)

Figure 100 PFS Ouwens model HR plot – random-effect model (Weibull)

Figure 101 PFS Ouwens model survival curve – random-effect model (log-normal)

Figure 102 PFS Ouwens model HR plot – random-effect model (log-normal)

Table 42 PFS Ouwens model monthly HRs – random-effect model BEV+IFN

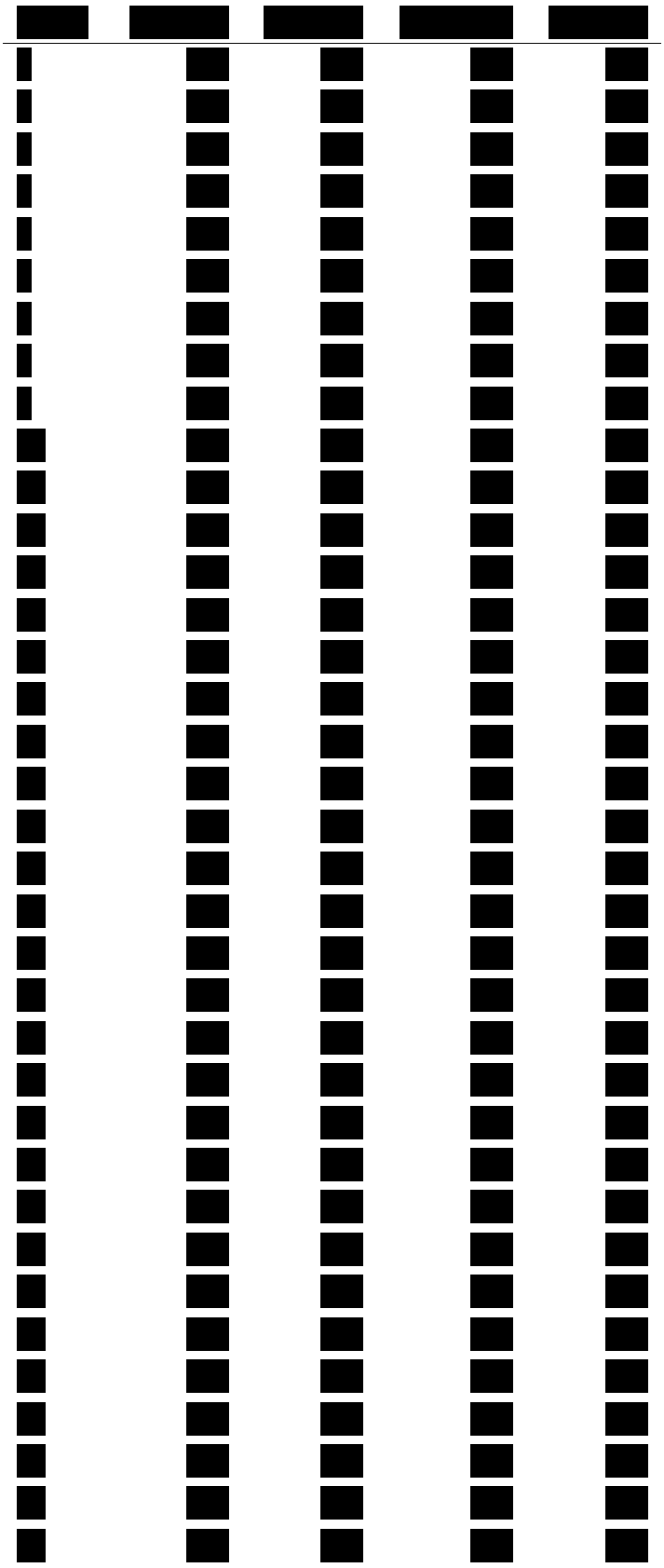


Table 43 PFS Ouwens model monthly HRs – random-effect model IFN

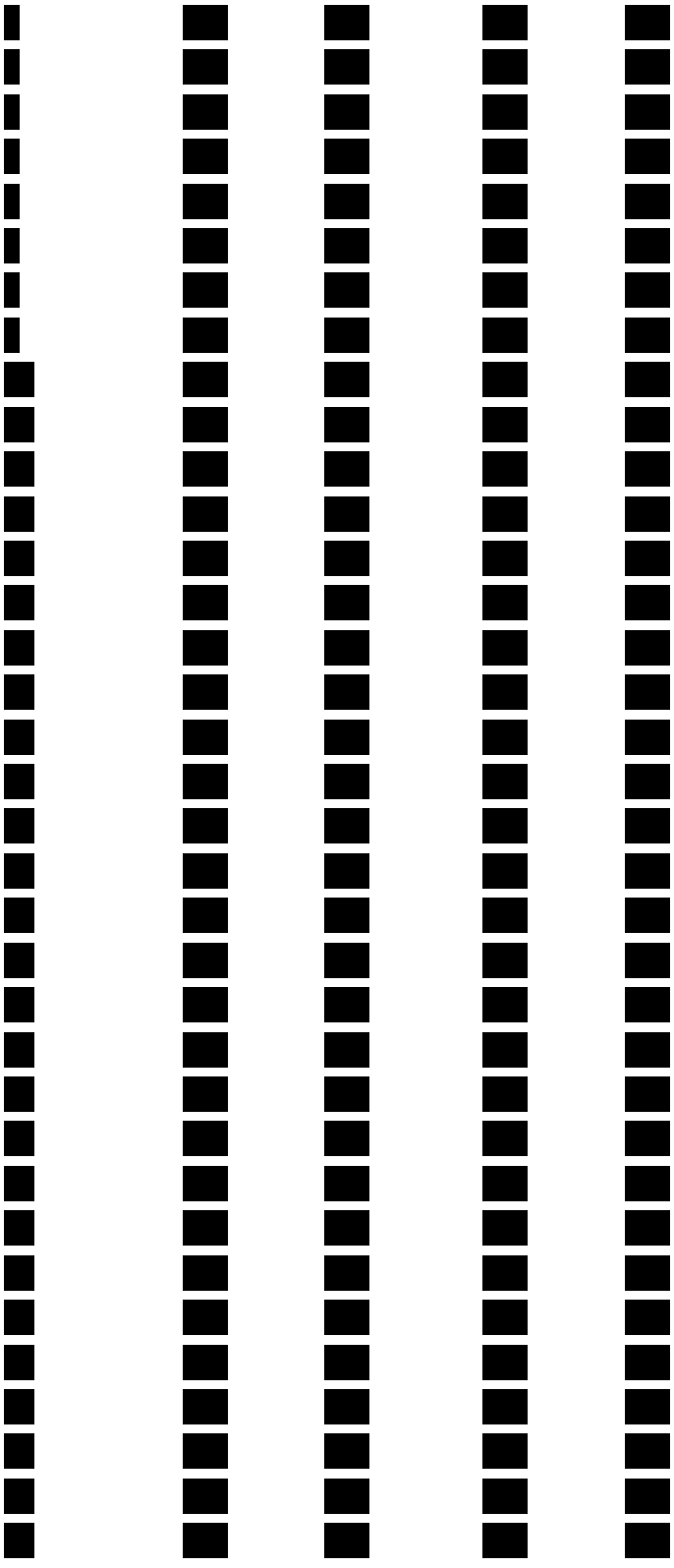
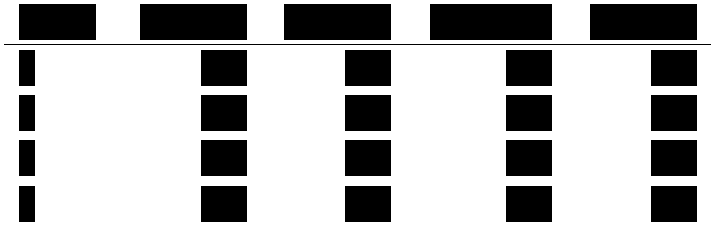


Table 44 PFS Ouwens model monthly HRs – random-effect model pazopanib



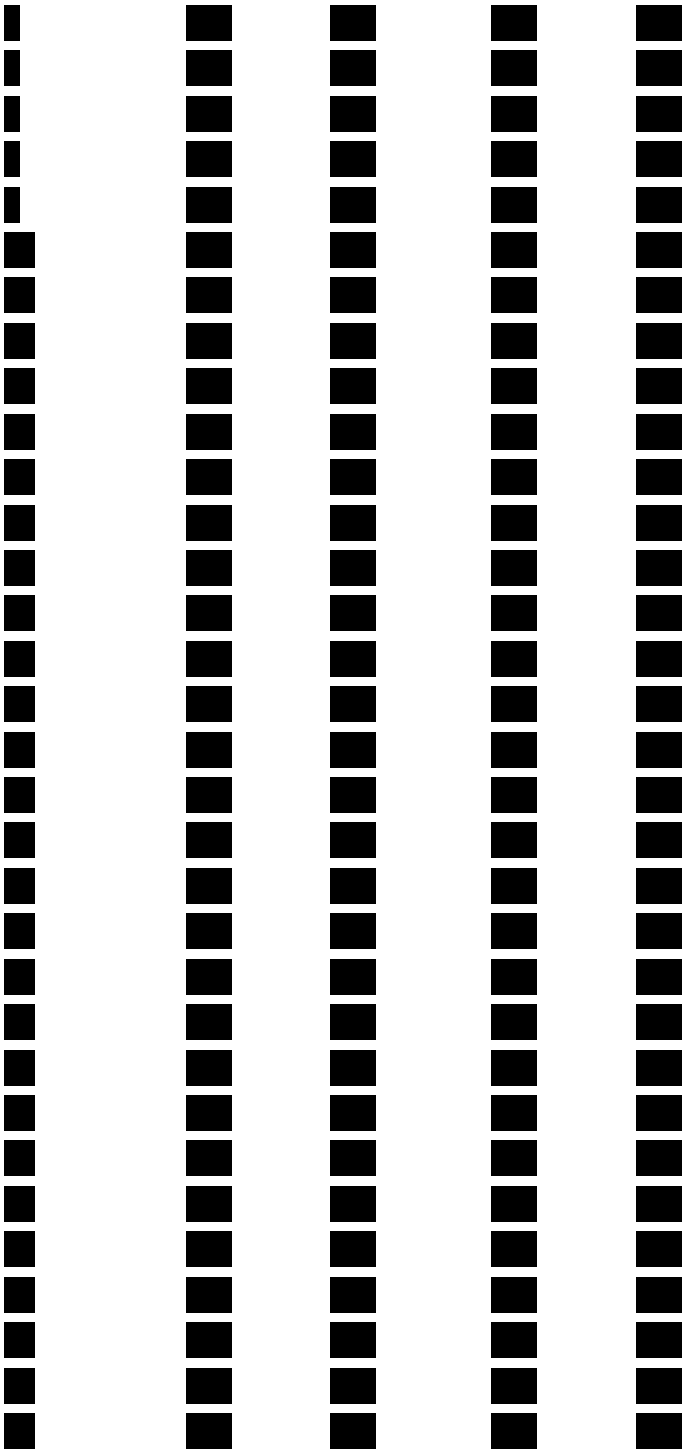


Table 45 PFS Ouwens model monthly HRs – random-effect model placebo

[redacted]

Table 46 PFS Ouwens model monthly HRs – random-effect model sorafenib

[redacted]

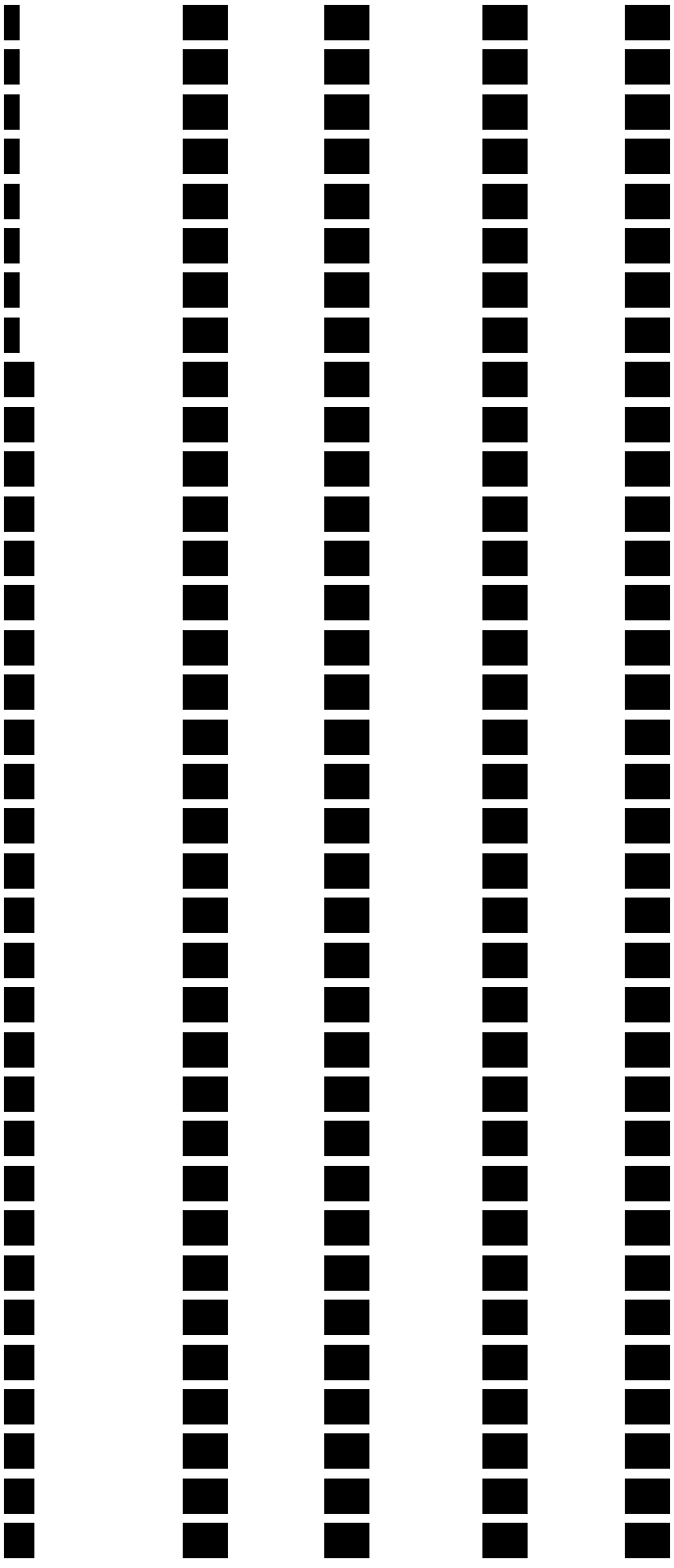
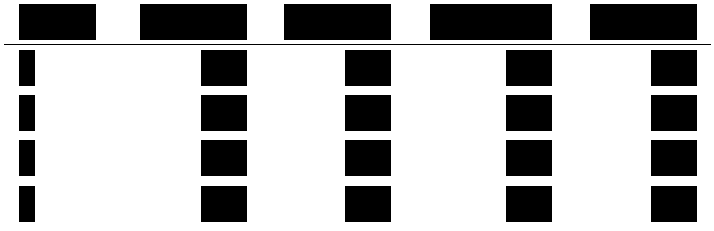


Table 47 PFS Ouwens model monthly HRs – random-effect model sunitinib



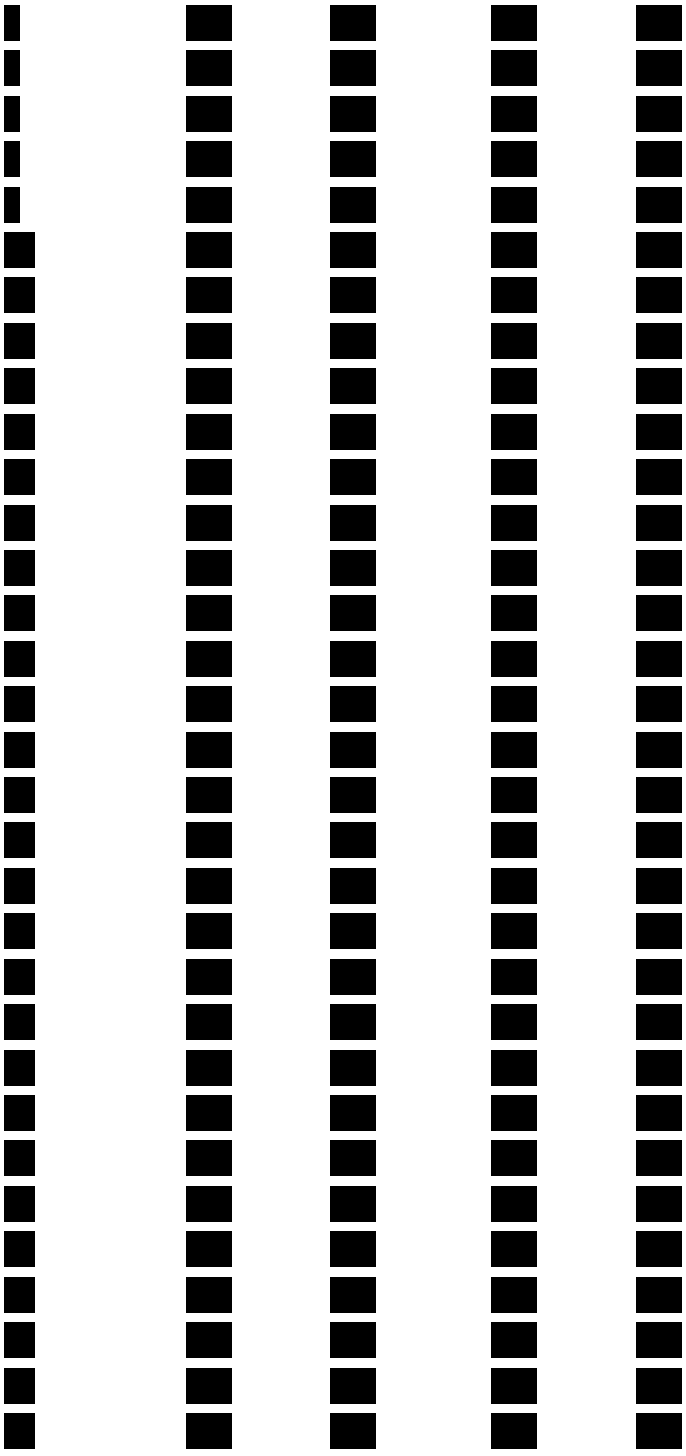
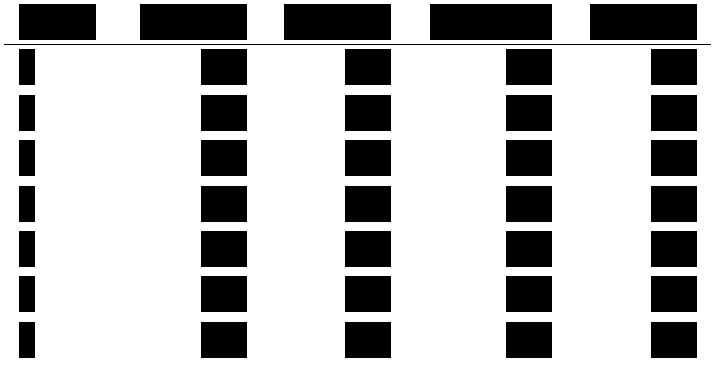


Table 48 PFS Ouwens model monthly HRs – random-effect model temsirolimus



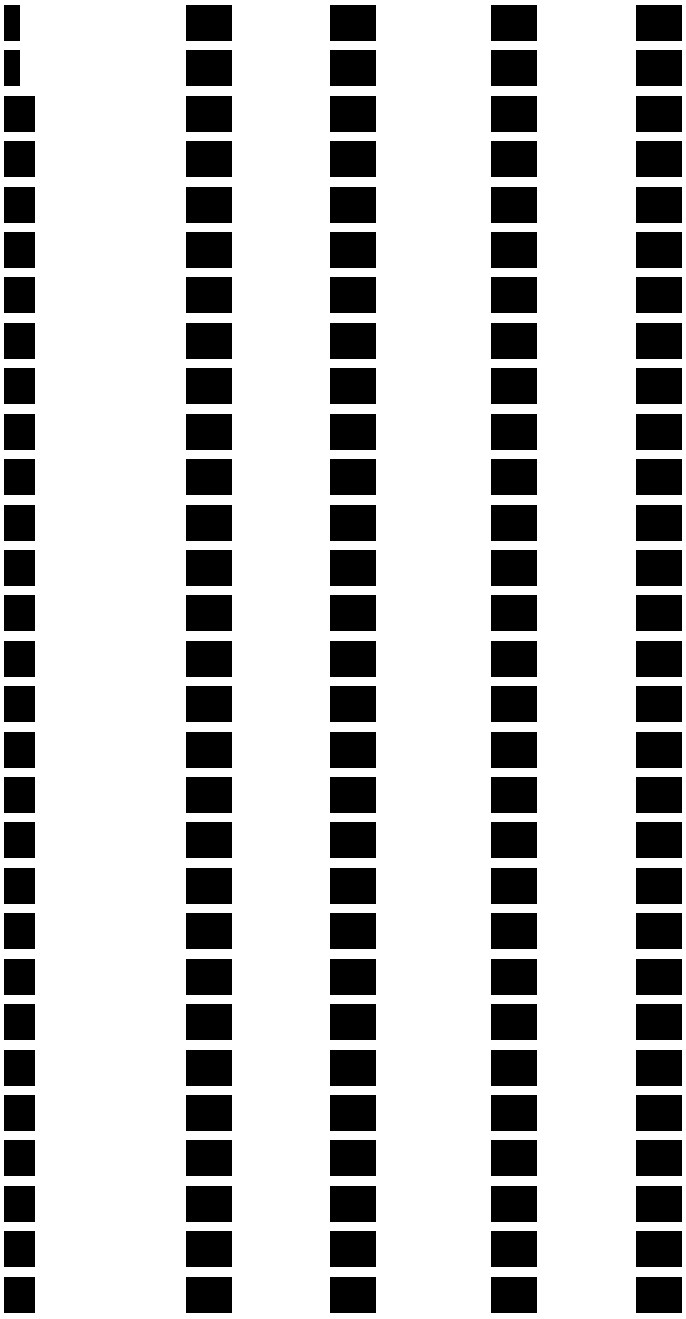
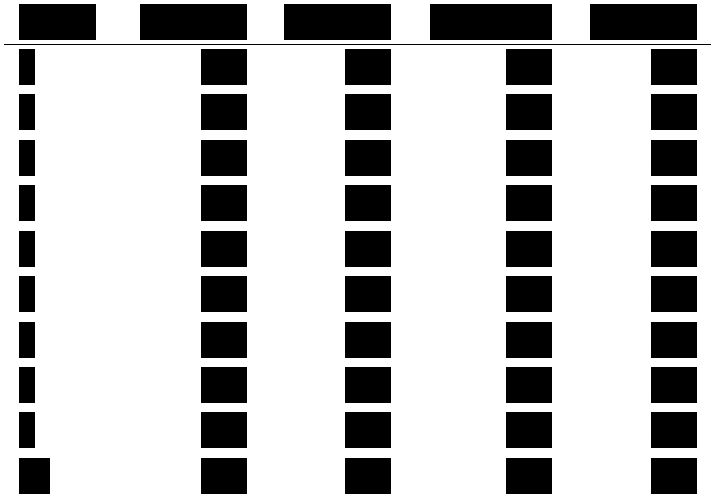


Table 49 PFS Ouwens model monthly HRs – random-effect model tivozanib



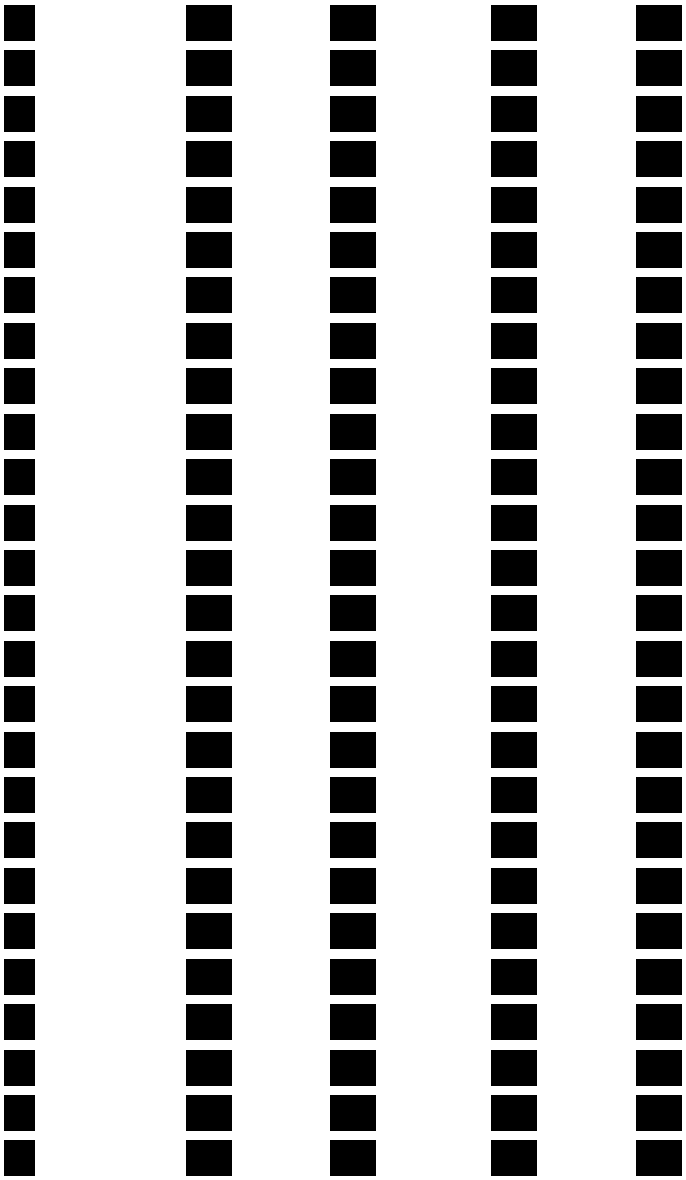


Table 50 Statistical fits – DIC Ouwens random-effect

parameters	OS	PFS
log-logistic	5666.3	9203.4
exponential	5778.5	9643.1
Weibull	5756.1	9561.4
lognormal	5615.4	9074.1

Fractional polynomial based on fixed effect model

OS fractional polynomial 1st order

Figure 103 OS fractional polynomial model survival curve – fixed-effect model (P=0)

Figure 104 OS fractional polynomial model HR plot – fixed-effect model (P=0)

Figure 105 OS fractional polynomial model survival curve – fixed-effect model (P=1)

Figure 106 OS fractional polynomial model HR plot – fixed-effect model (P=1)

Figure 107 OS fractional polynomial model survival curve – fixed-effect model (P=0.5)

Figure 108 OS fractional polynomial model HR plot – fixed-effect model (P=0.5)

Figure 109 OS fractional polynomial model survival curve – fixed-effect model (P=-1)

Figure 110 OS fractional polynomial model HR plot – fixed-effect model (P=-1)

Figure 111 OS fractional polynomial model survival curve – fixed-effect model (P=-0.5)

Figure 112 OS fractional polynomial model HR plot – fixed-effect model (P=-0.5)

Table 51 OS fractional polynomial model monthly HRs – random-effect model BEV+IFN

[redacted]

Table 52 OS fractional polynomial model monthly HRs – random-effect model IFN

[redacted]

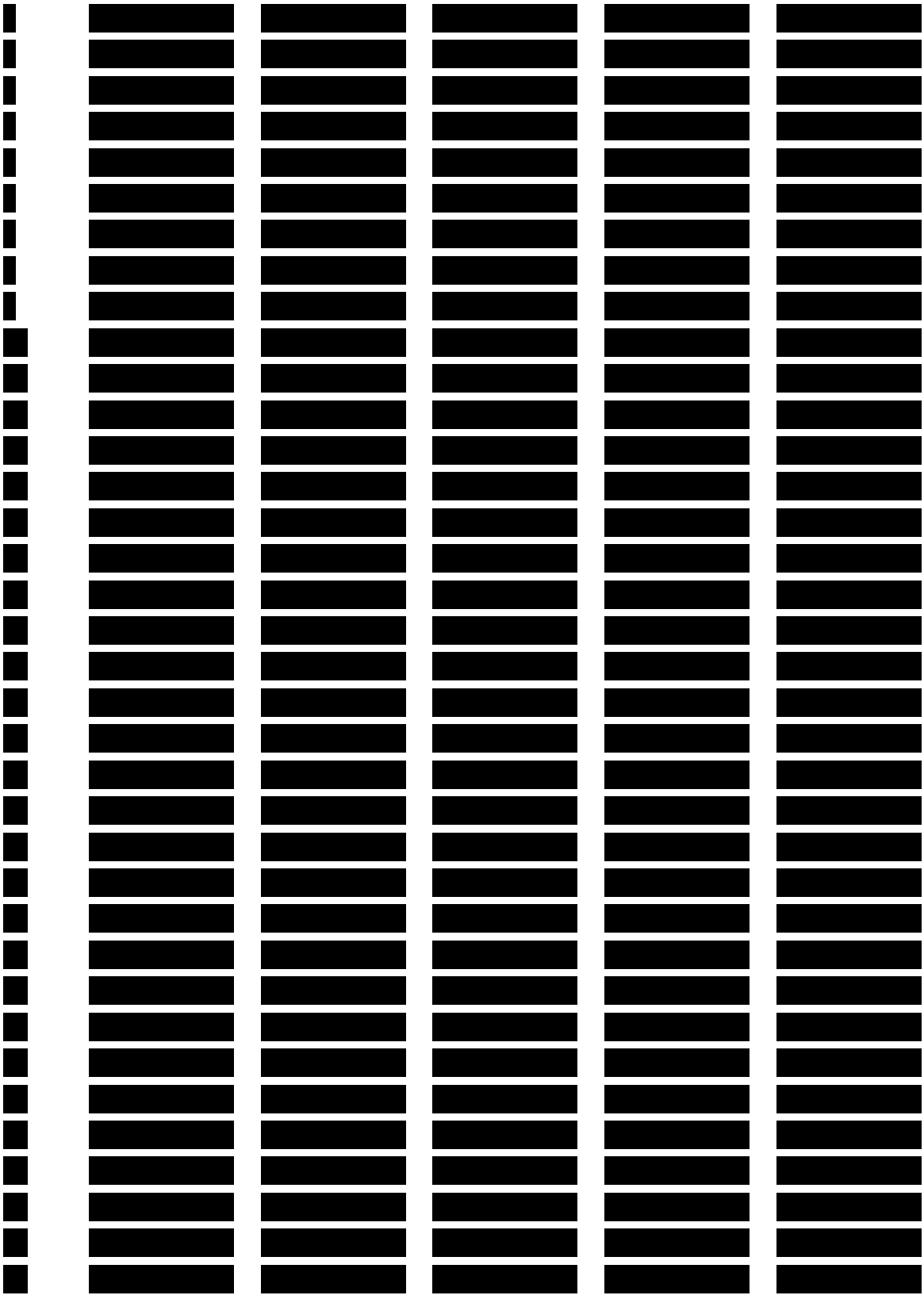
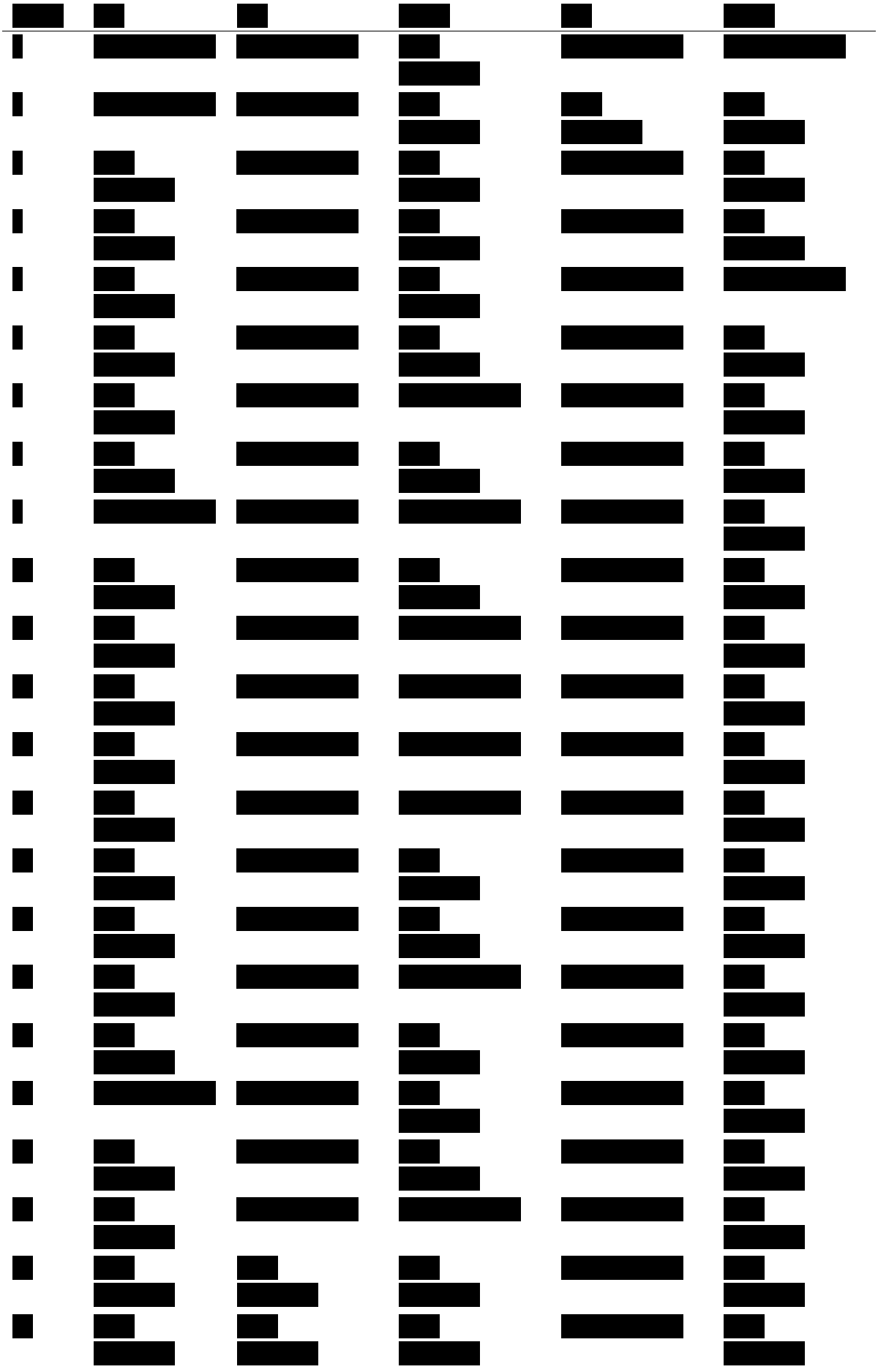


Table 53 OS fractional polynomial model monthly HRs – random-effect model pazopanib

[redacted]

Table 54 OS fractional polynomial model monthly HRs – random-effect model sorafenib



[redacted]

Table 55 OS fractional polynomial model monthly HRs – random-effect model sunitinib

[redacted]

[redacted]

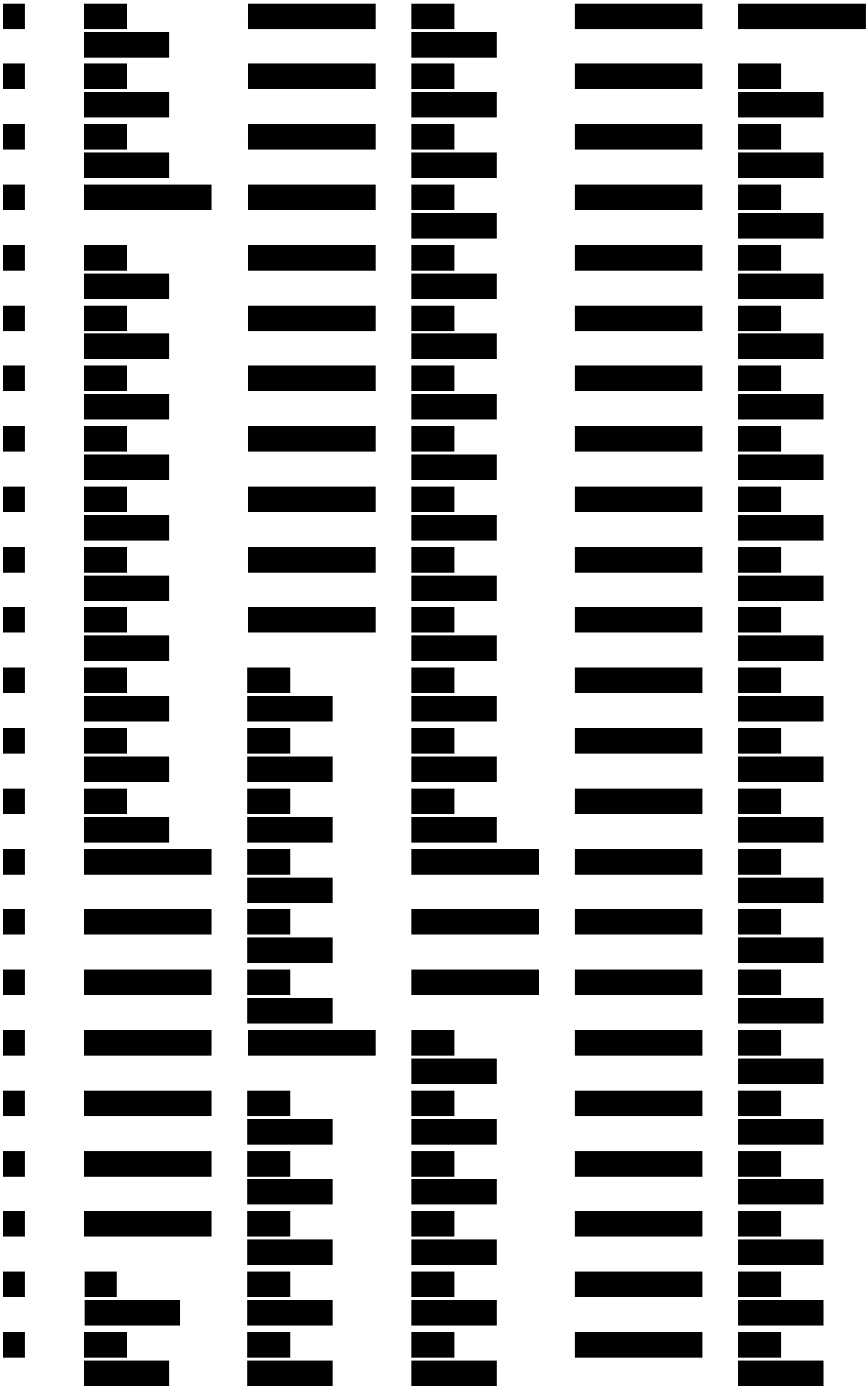
Table 56 OS fractional polynomial model monthly HRs – random-effect model temsirolimus

[redacted]

[redacted]

Table 57 OS fractional polynomial model monthly HRs – random-effect model tivozanib

[redacted]



[redacted]

OS fractional polynomial 2nd order

Figure 113 OS fractional polynomial model survival curve – fixed-effect model (P1=-0.5, P2=0)

Figure 114 OS fractional polynomial model HR plot – fixed-effect model (P1=-0.5, P2=0)

Figure 115 OS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=0)

Figure 116 OS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=0)

Figure 117 OS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=-1)

Figure 118 OS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=-1)

Figure 119 OS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=0.5)

Figure 120 OS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=0.5)

Figure 121 OS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=1)

Figure 122 OS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=1)

Table 58 OS fractional polynomial model monthly HRs– fixed-effect model BEV+IFN

[redacted]

[redacted]

Table 59 OS fractional polynomial model monthly HRs– fixed-effect model IFN

[redacted]

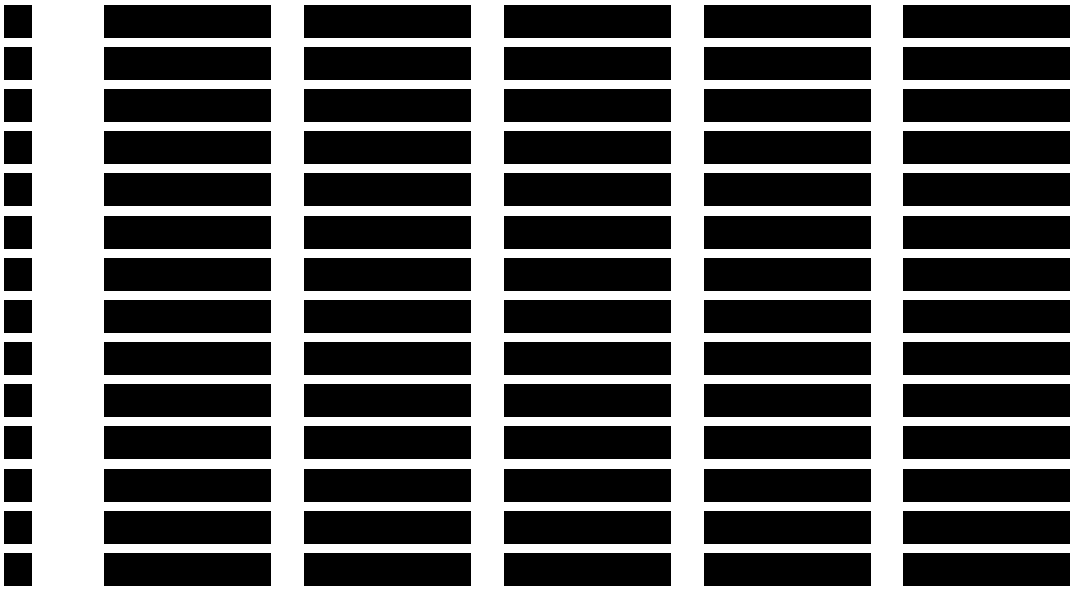


Table 60 OS fractional polynomial model monthly HRs– fixed-effect model pazopanib

Table 61 OS fractional polynomial model monthly HRs– fixed-effect model sorafenib

[redacted]

Table 62 OS fractional polynomial model monthly HRs– fixed-effect model sunitinib

[redacted]

[redacted]

Table 63 OS fractional polynomial model monthly HRs– fixed-effect model temsirolimus

[redacted]

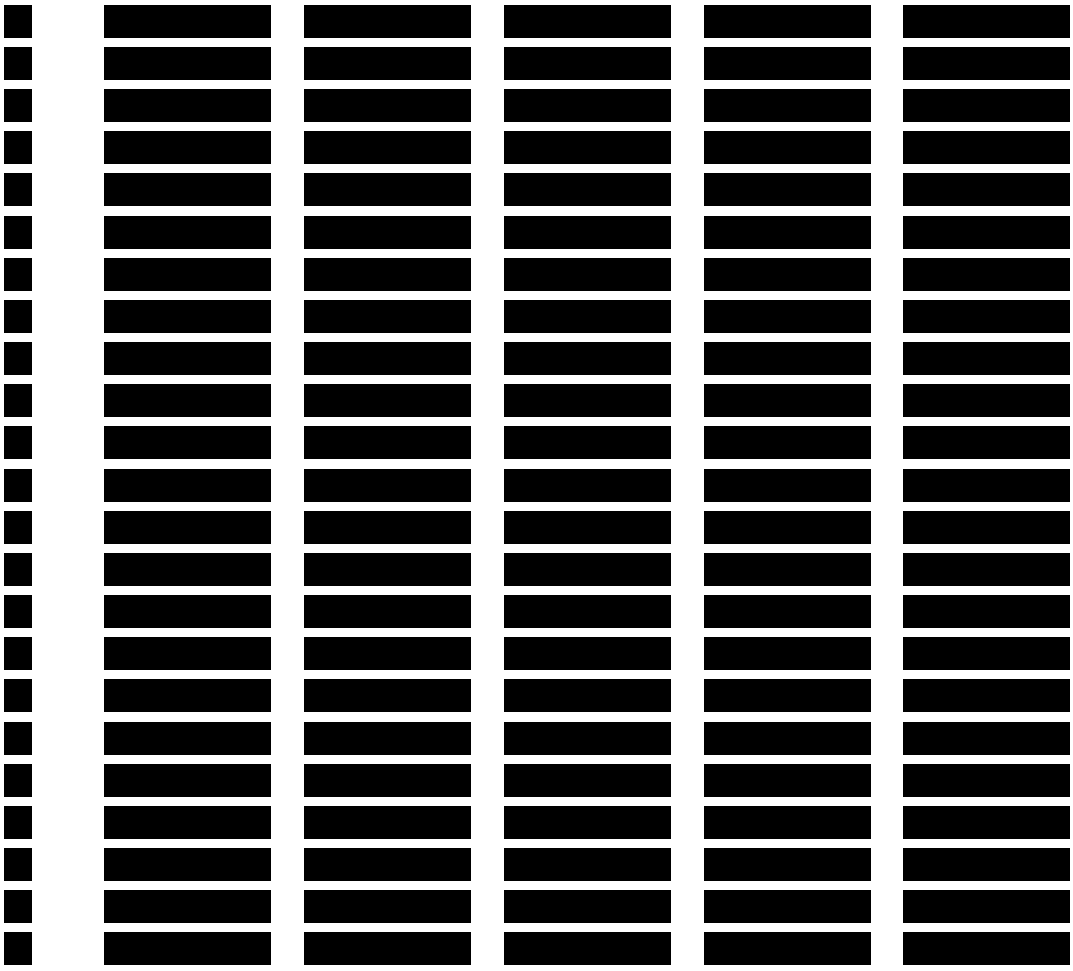
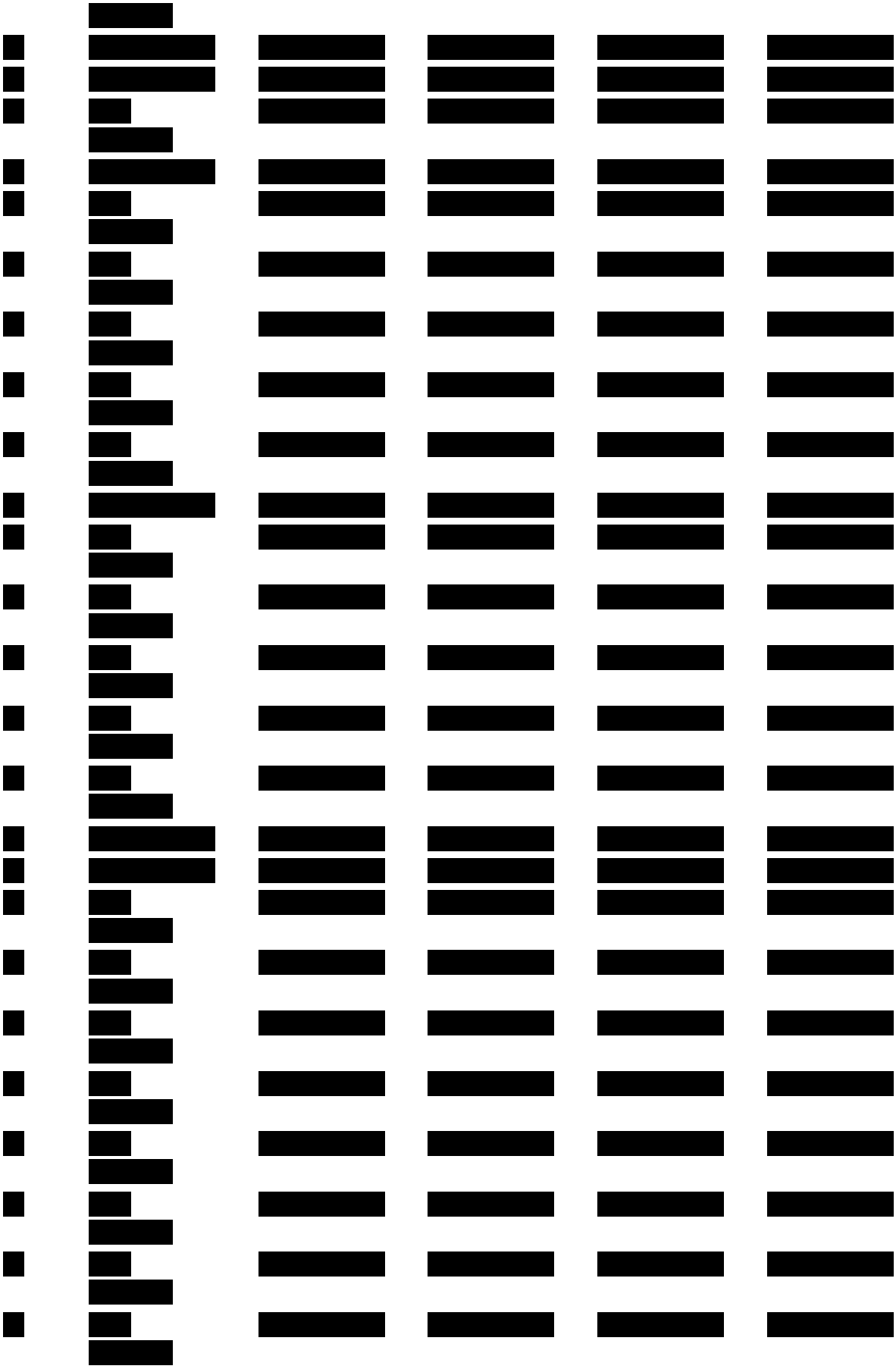


Table 64 OS fractional polynomial model monthly HRs– fixed-effect model tivozanib



PFS fractional polynomial 1st order

Figure 123 PFS fractional polynomial model survival curve – fixed-effect model (P=0)

Figure 124 PFS fractional polynomial model HR plot – fixed-effect model (P=0)

Figure 125 PFS fractional polynomial model survival curve – fixed-effect model (P=1)

Figure 126 PFS fractional polynomial model HR plot – fixed-effect model (P=1)

Figure 127 PFS fractional polynomial model survival curve – fixed-effect model (P=0.5)

Figure 128 PFS fractional polynomial model HR plot – fixed-effect model (P=0.5)

Figure 129 PFS fractional polynomial model survival curve – fixed-effect model (P=-1)

Figure 130 PFS fractional polynomial model HR plot – fixed-effect model (P=-1)

Figure 131 PFS fractional polynomial model survival curve – fixed-effect model (P=-0.5)

Figure 132 PFS fractional polynomial model HR plot – fixed-effect model (P=-0.5)

Table 65 PFS fractional polynomial model monthly HRs– fixed-effect model BEV+IFN

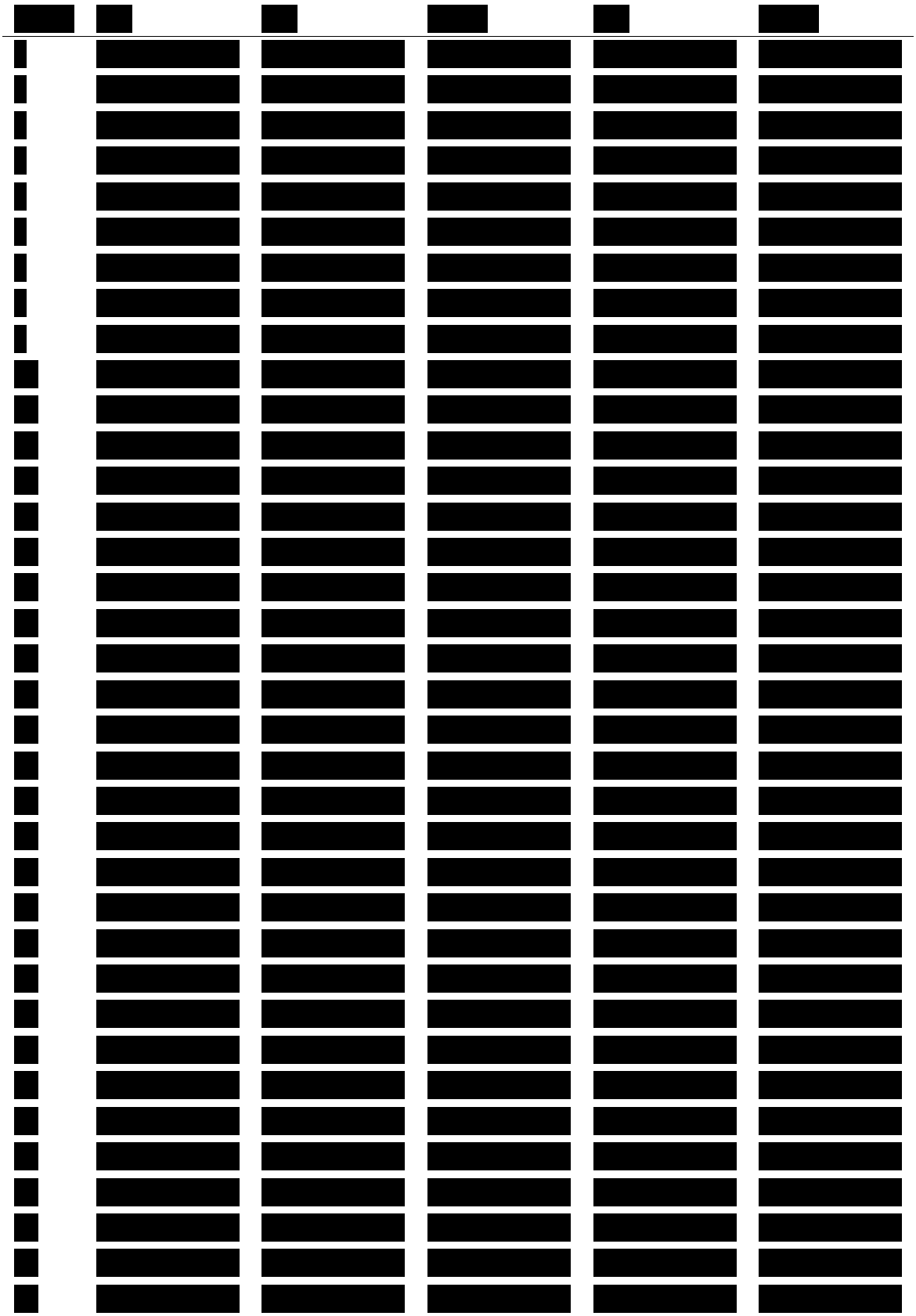


Table 66 PFS fractional polynomial model monthly HRs– fixed-effect model IFN

Table 67 PFS fractional polynomial model monthly HRs– fixed-effect model pazopanib

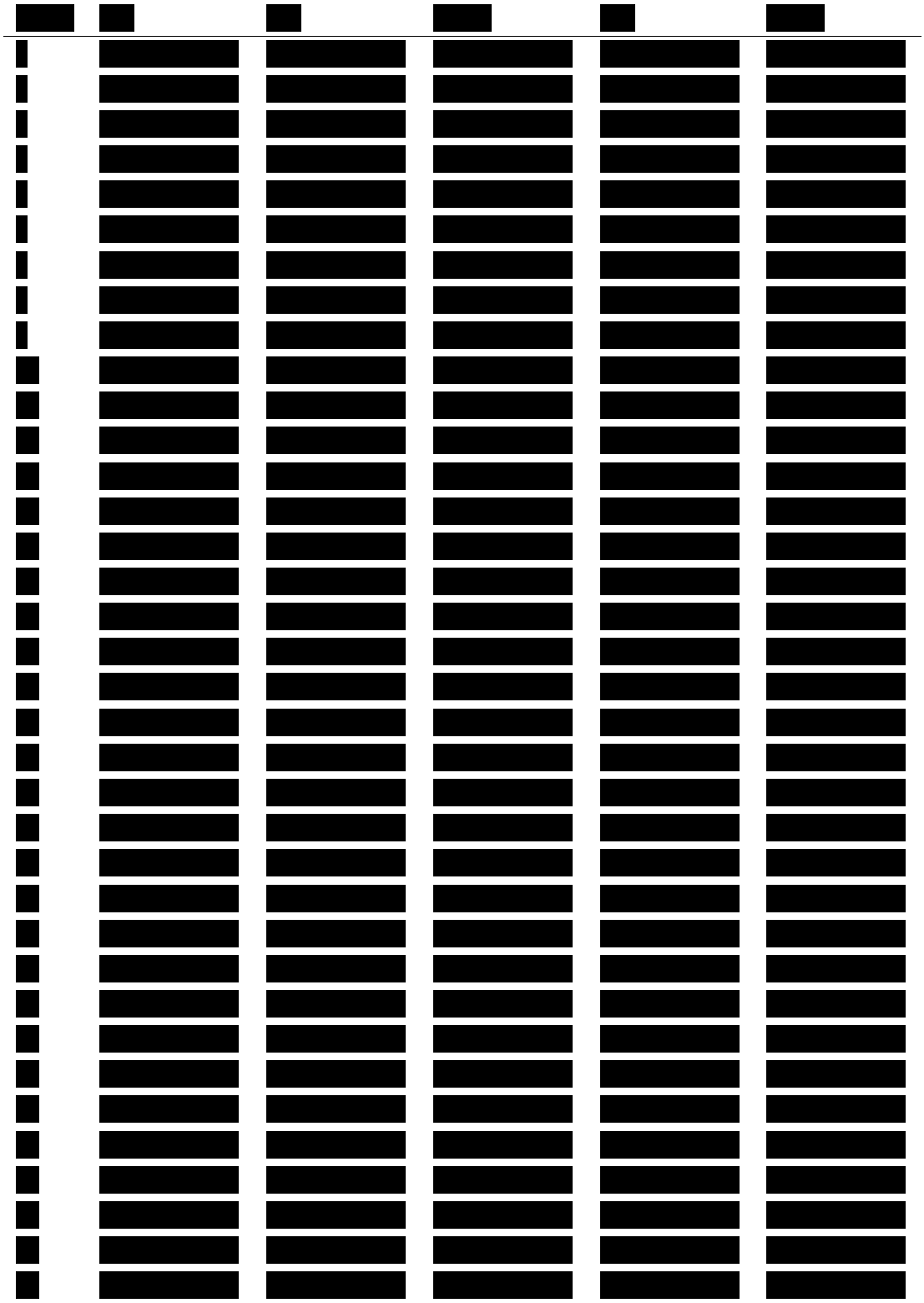


Table 68 PFS fractional polynomial model monthly HRs– fixed-effect model placebo

[redacted]

Table 69 PFS fractional polynomial model monthly HRs– fixed-effect model sorafenib

[redacted]

Table 70 PFS fractional polynomial model monthly HRs– fixed-effect model sunitinib

[redacted]

Table 71 PFS fractional polynomial model monthly HRs– fixed-effect model temsirolimus

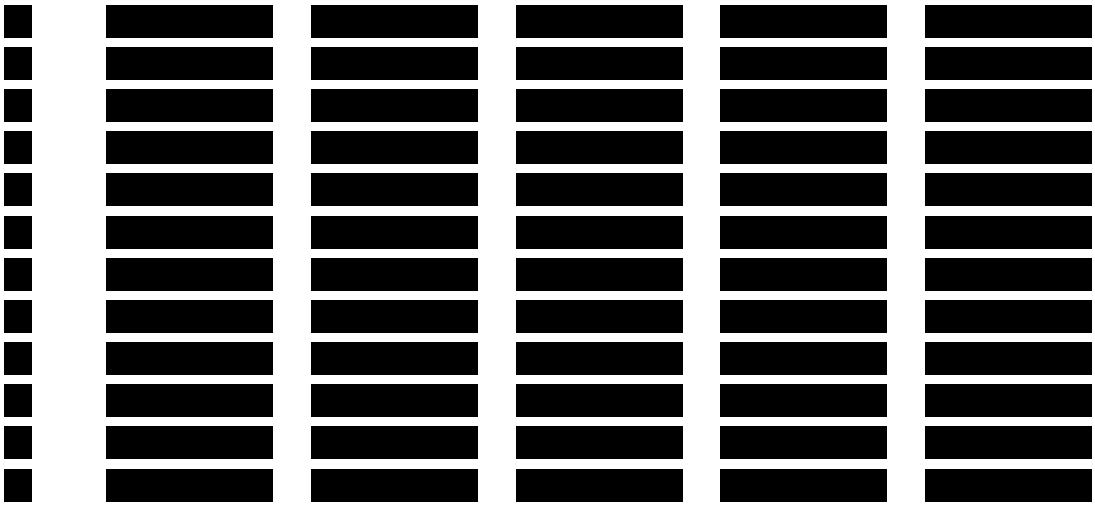


Table 72 PFS fractional polynomial model monthly HRs– fixed-effect model tivozanib

[redacted]

PFS fractional polynomial 2nd order

Figure 133 PFS fractional polynomial model survival curve – fixed-effect model (P1=-0.5, P2=0)

Figure 134 PFS fractional polynomial model HR plot – fixed-effect model (P1=-0.5, P2=0)

Figure 135 PFS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=0)

Figure 136 PFS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=0)

Figure 137 PFS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=-1)

Figure 138 PFS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=-1)

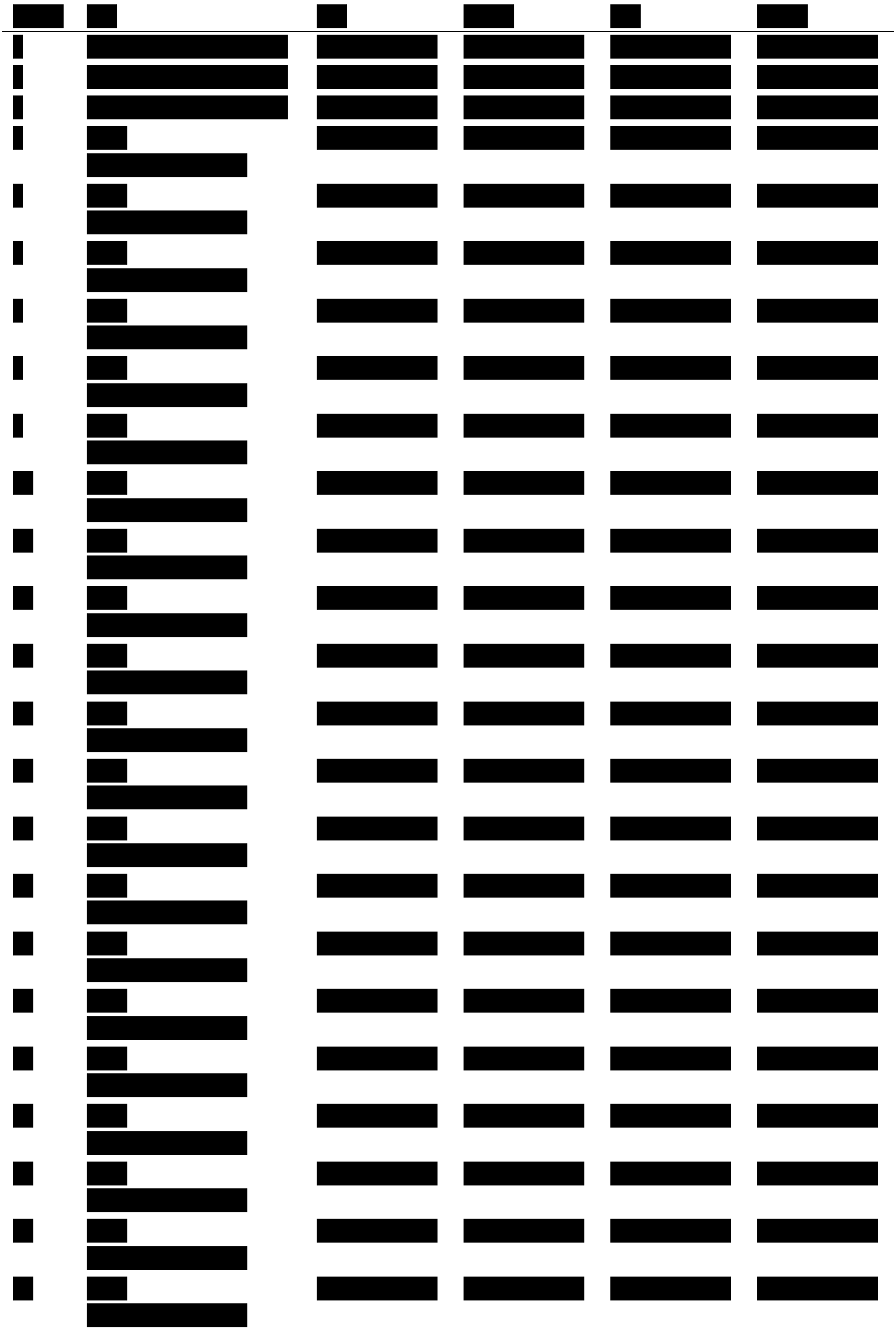
Figure 139 PFS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=0.5)

Figure 140 PFS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=0.5)

Figure 141 PFS fractional polynomial model survival curve – fixed-effect model (P1=-1, P2=1)

Figure 142 PFS fractional polynomial model HR plot – fixed-effect model (P1=-1, P2=1)

Table 73 PFS fractional polynomial model monthly HRs– fixed-effect model BEV+IFN



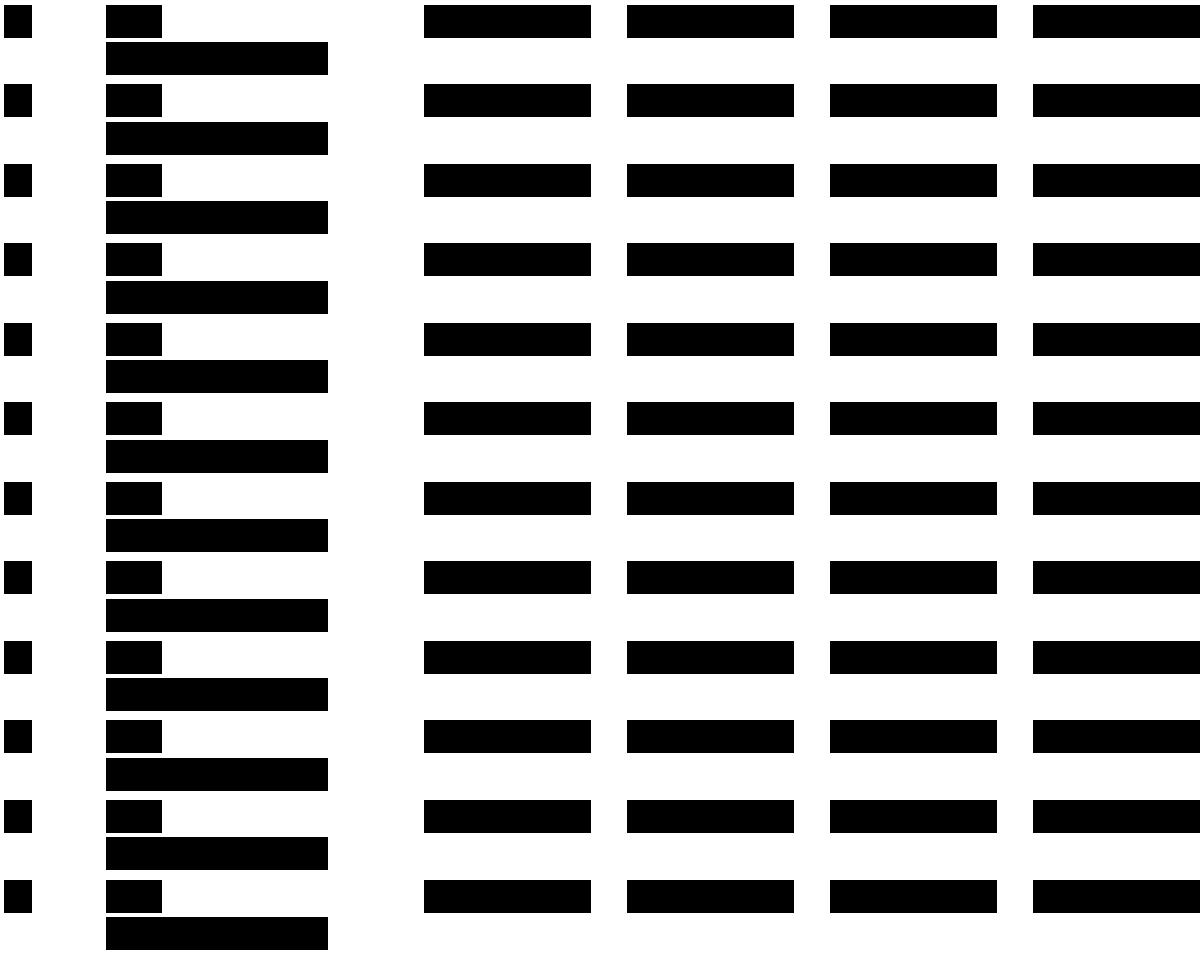
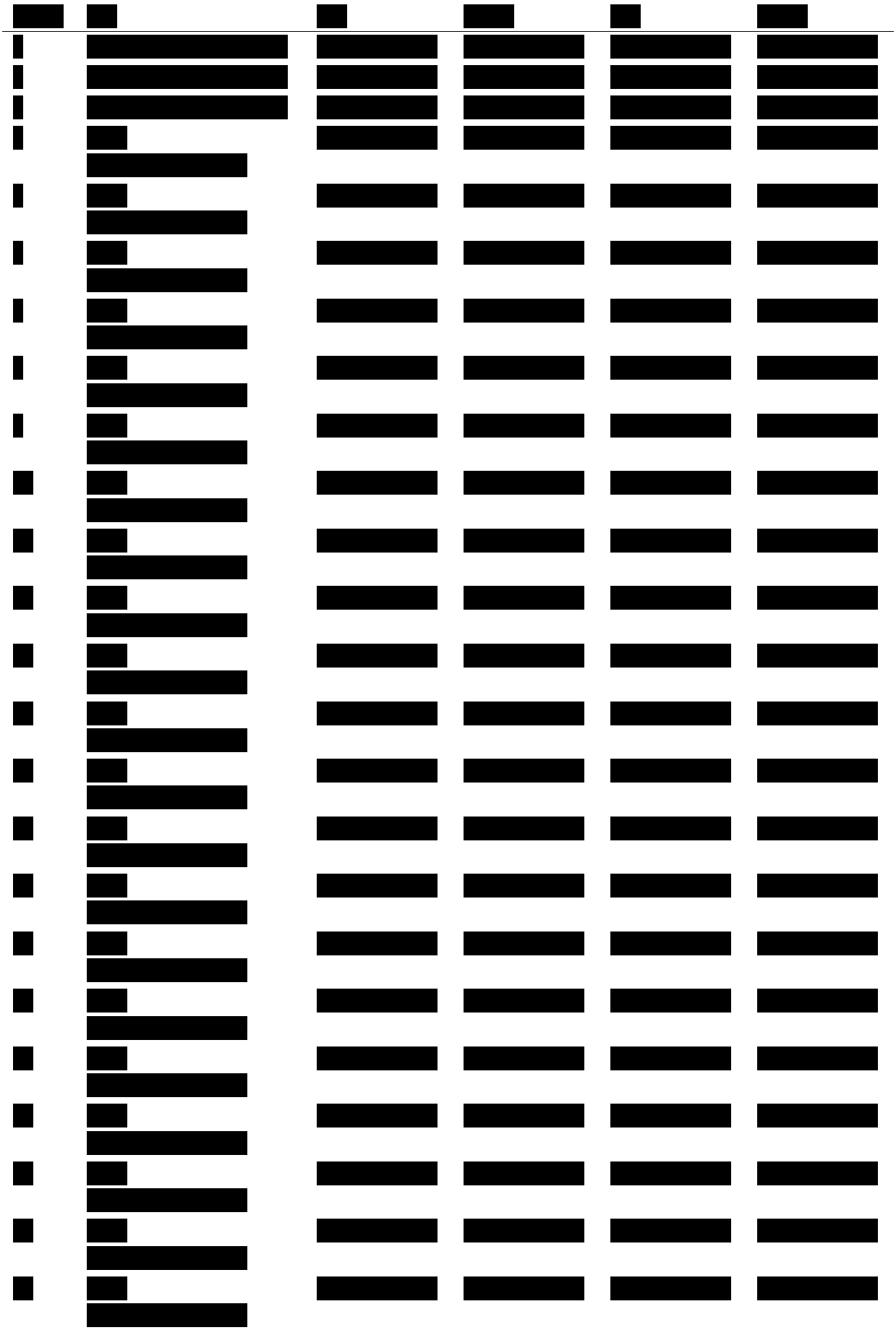


Table 74 PFS fractional polynomial model monthly HRs– fixed-effect model IFN



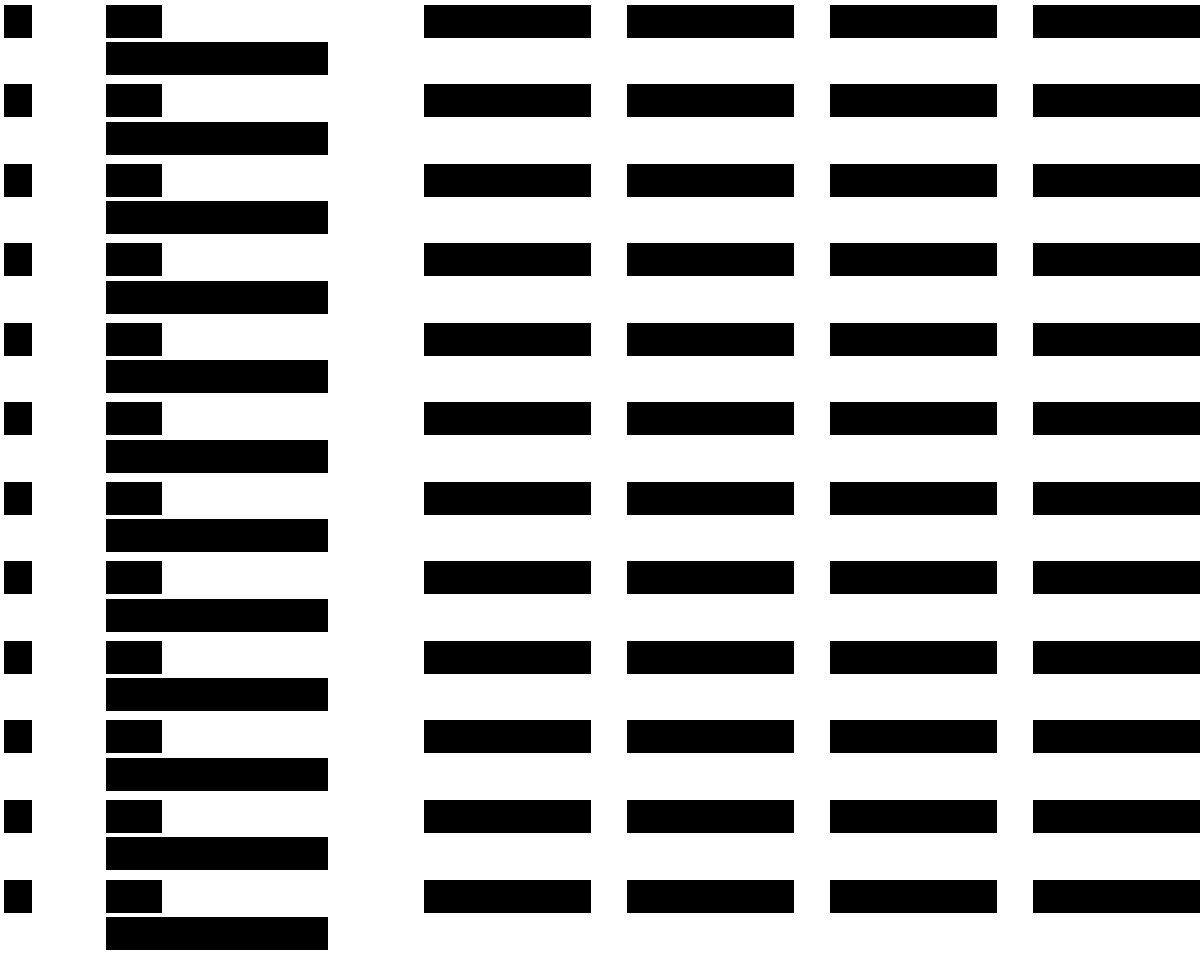
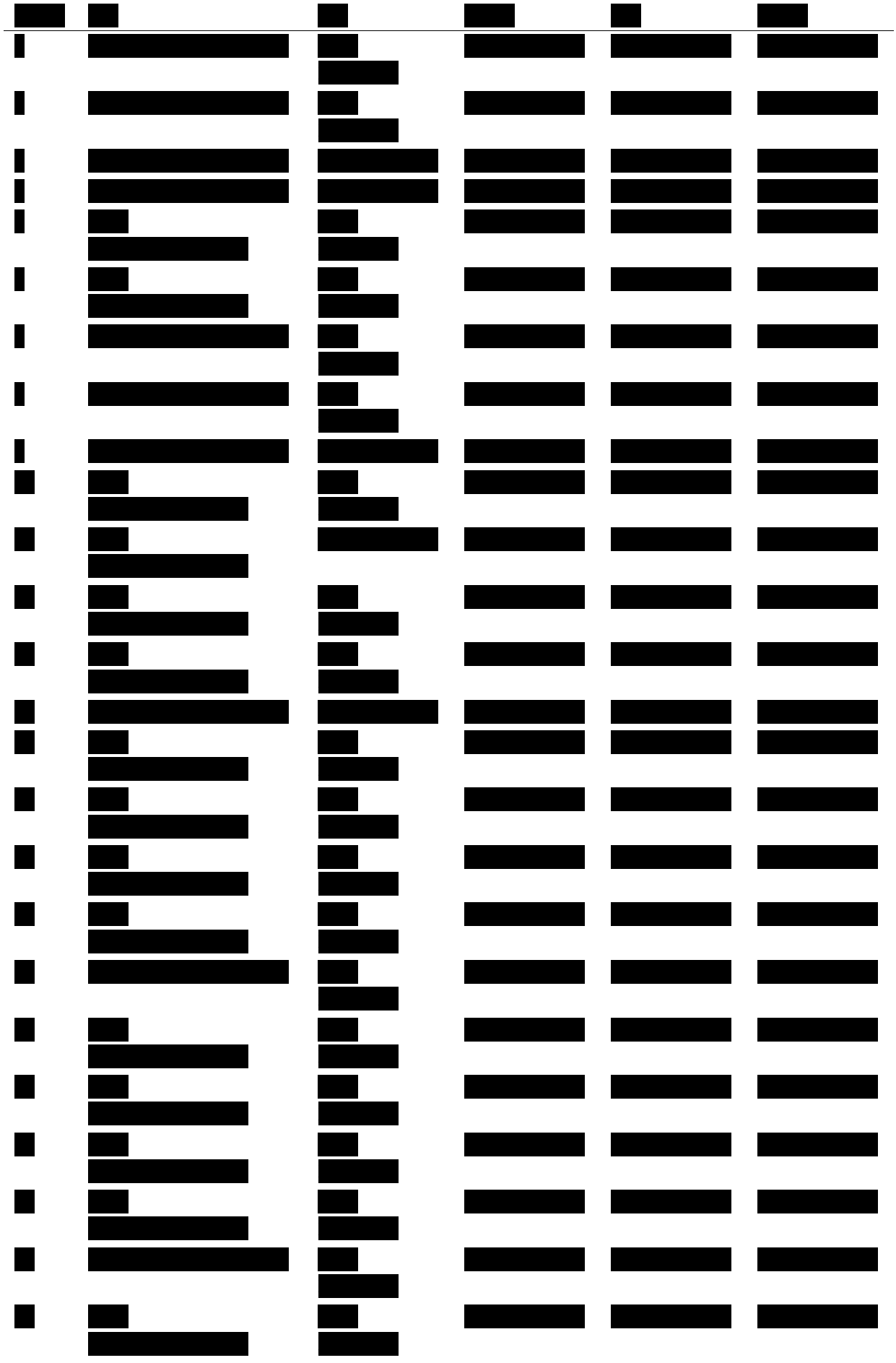


Table 75 PFS fractional polynomial model monthly HRs– fixed-effect model pazopanib



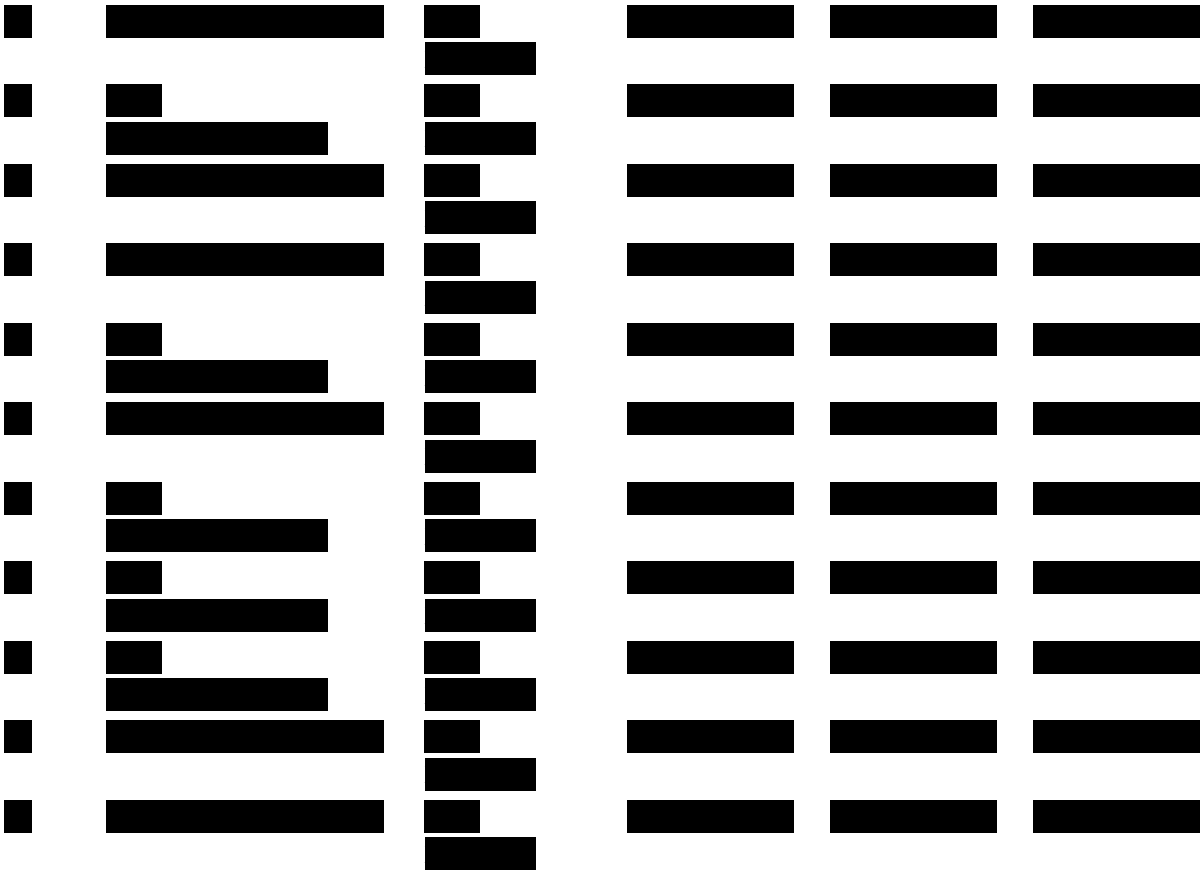


Table 76 PFS fractional polynomial model monthly HRs– fixed-effect model placebo

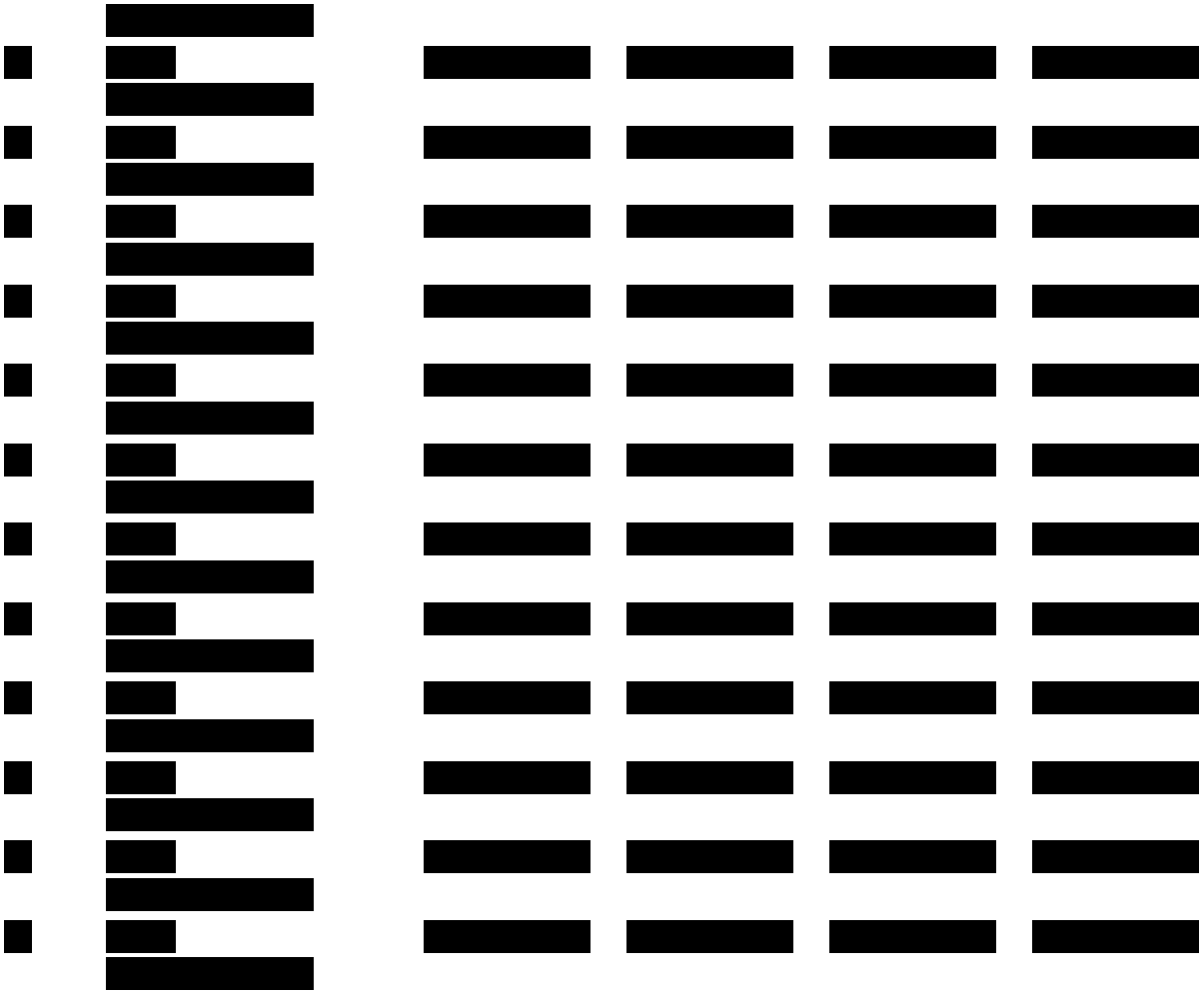


Table 77 PFS fractional polynomial model monthly HRs– fixed-effect model sorafenib

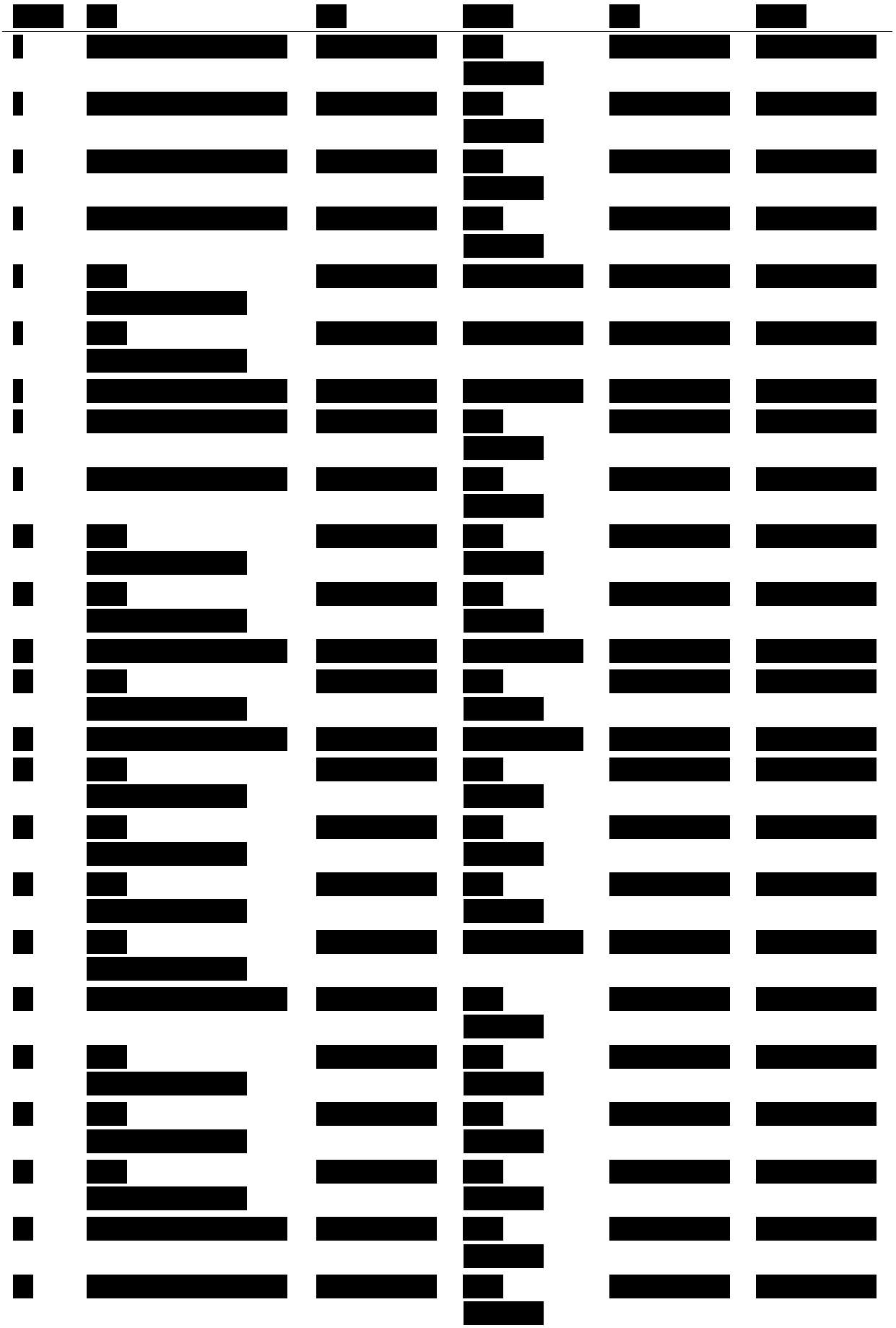
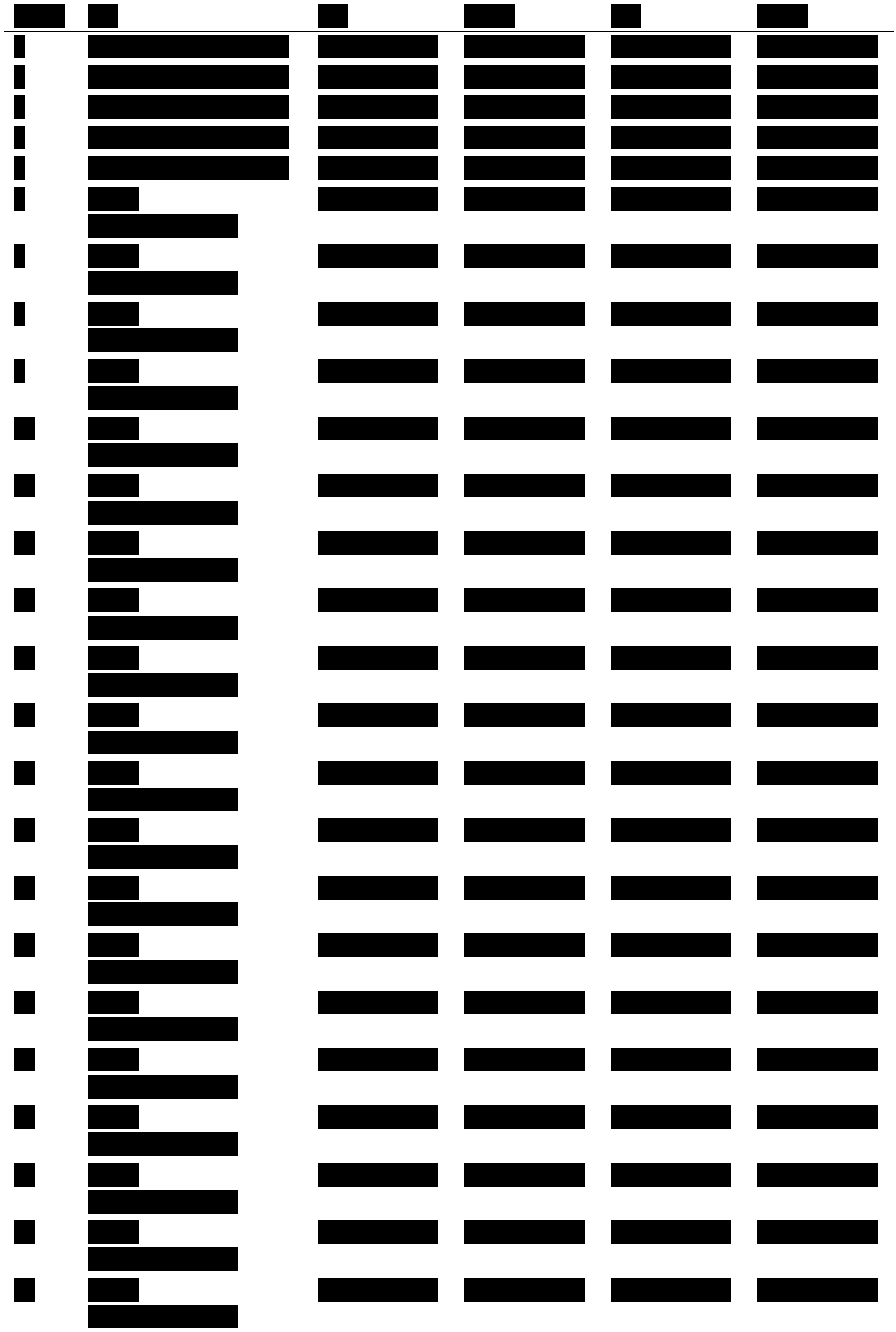


Table 78 PFS fractional polynomial model monthly HRs– fixed-effect model sunitinib



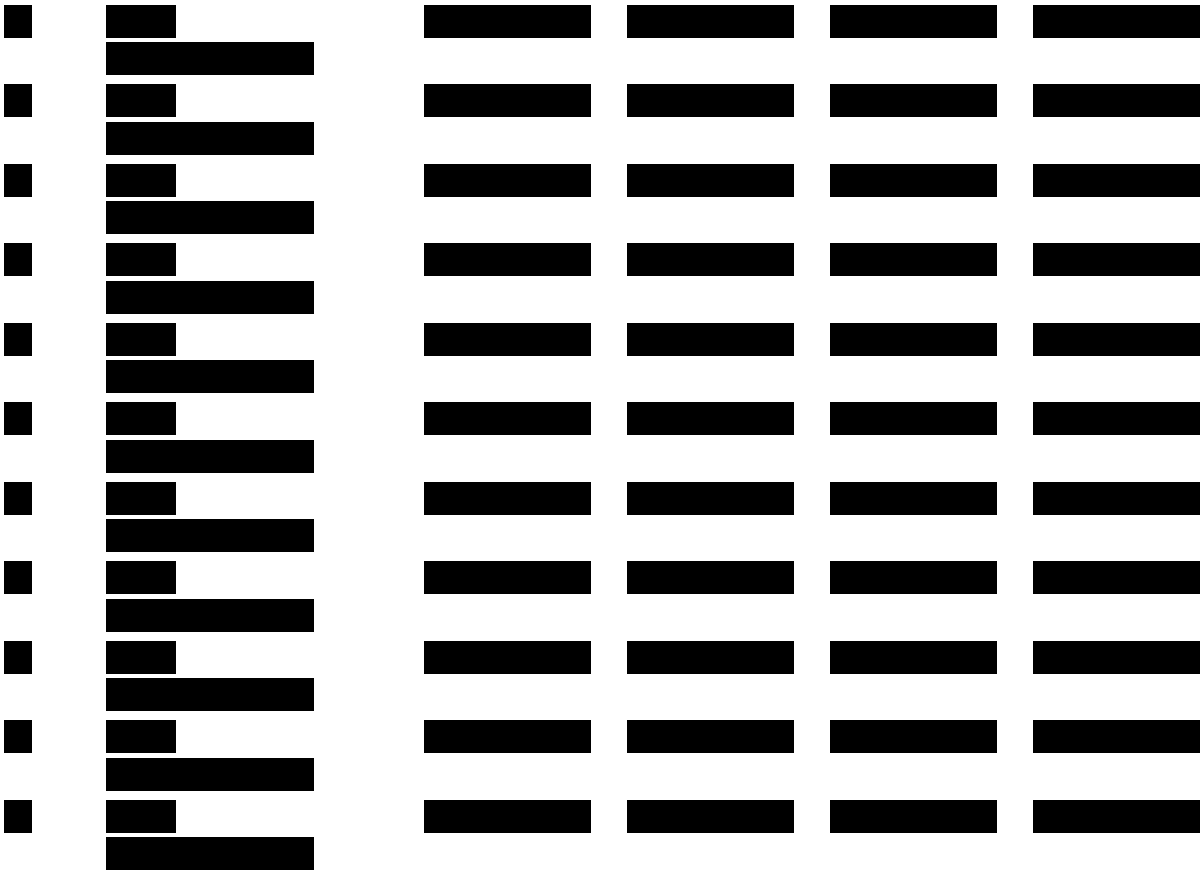
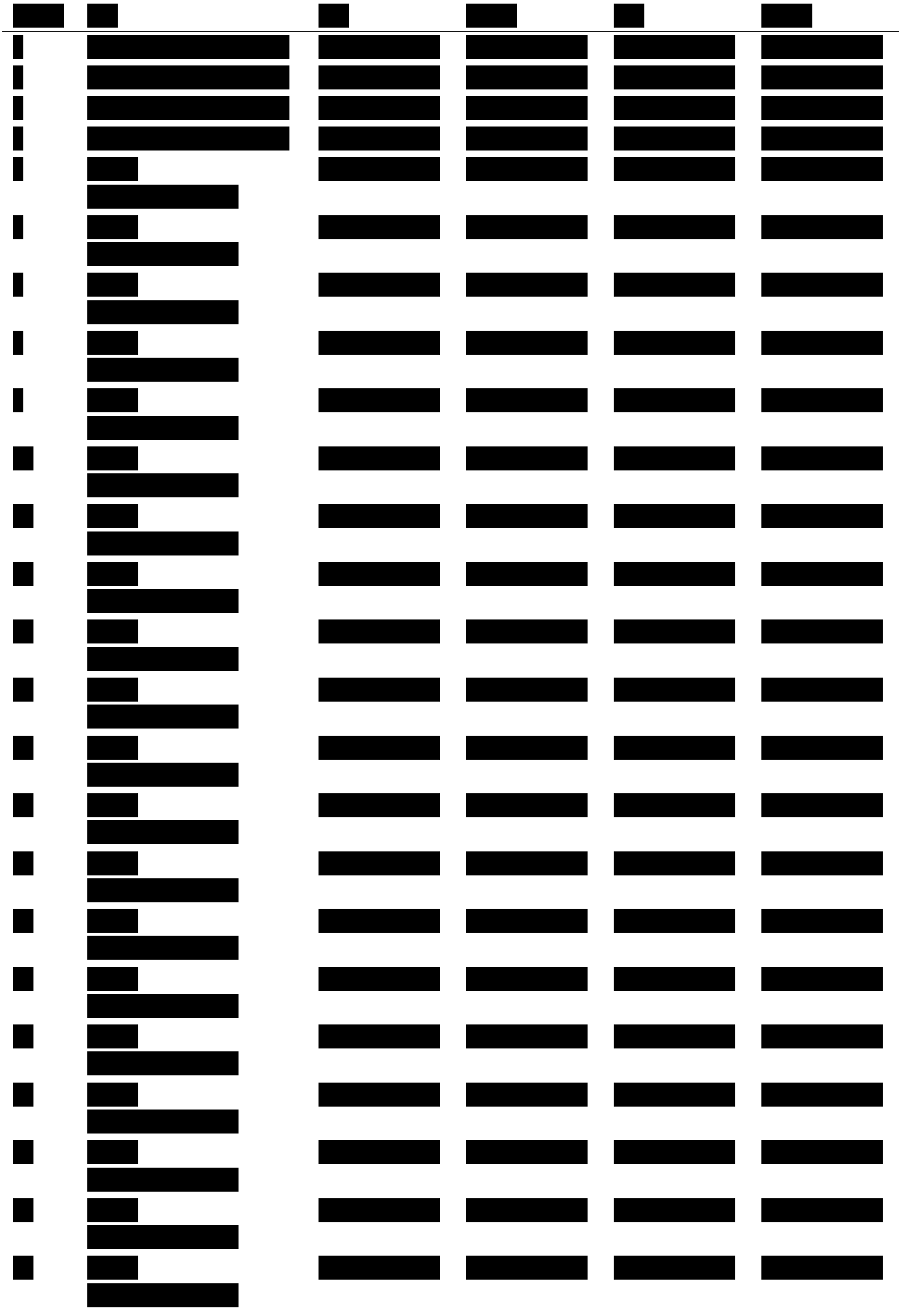


Table 79 PFS fractional polynomial model monthly HRs– fixed-effect model temsirolimus



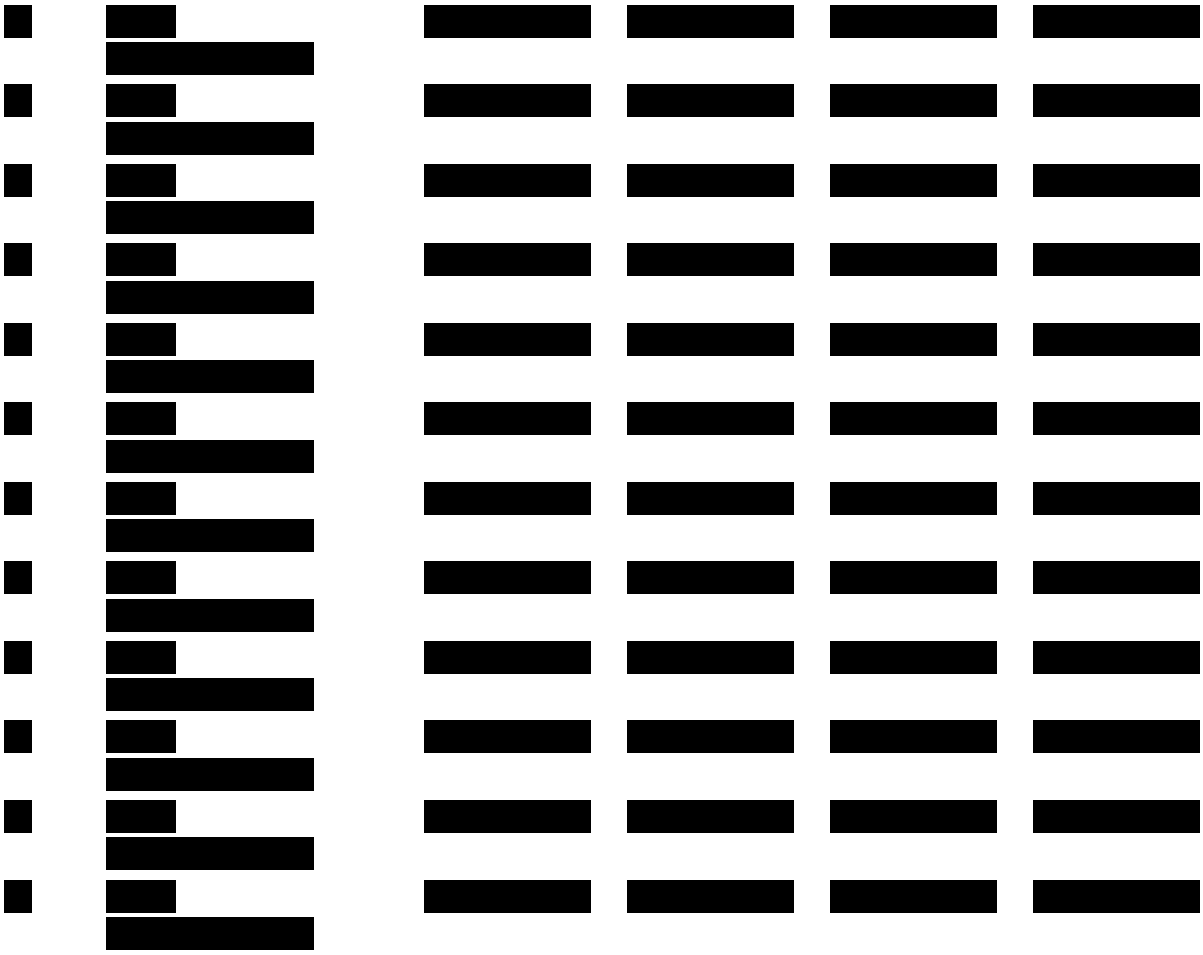
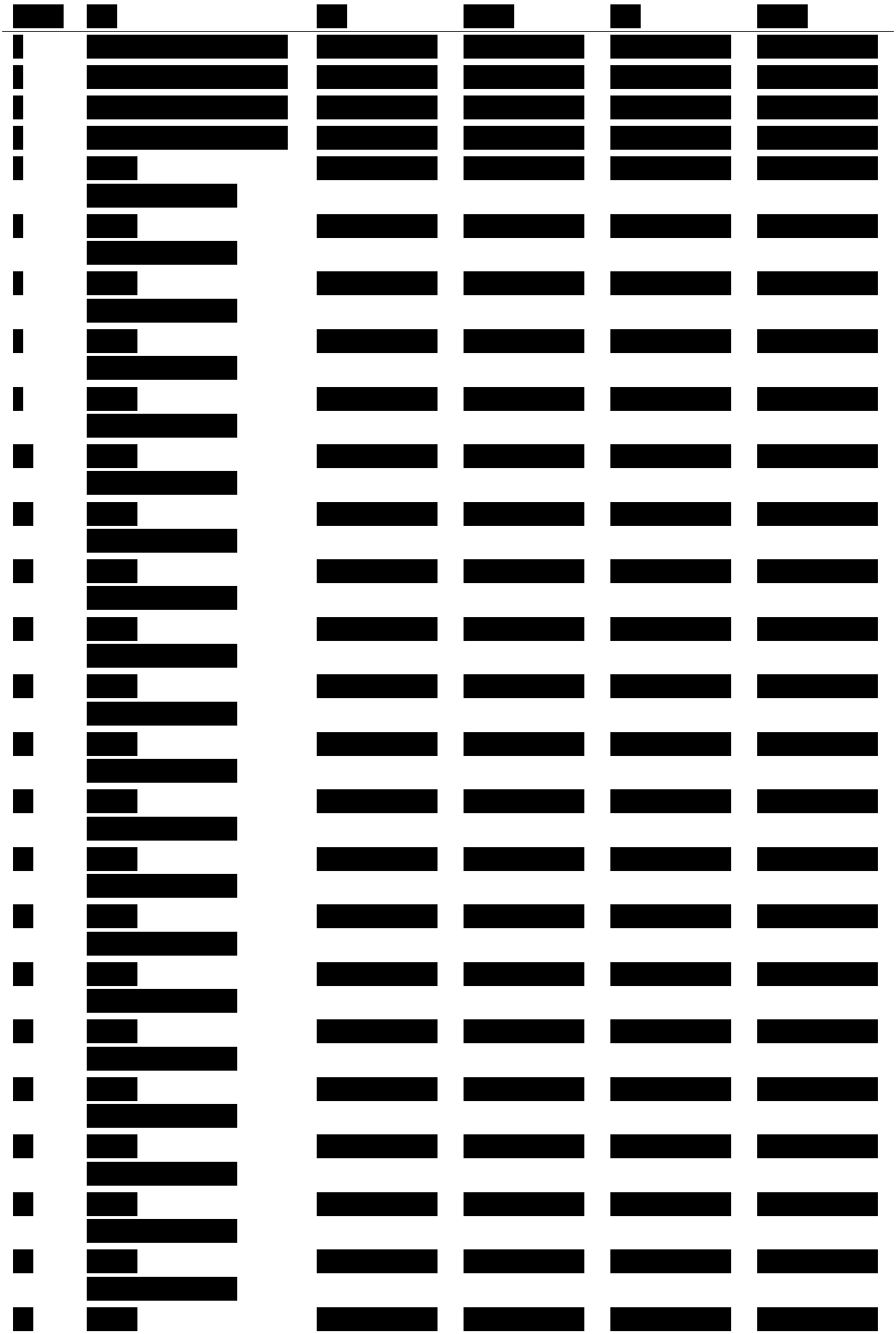


Table 80 PFS fractional polynomial model monthly HRs– fixed-effect model tivozanib



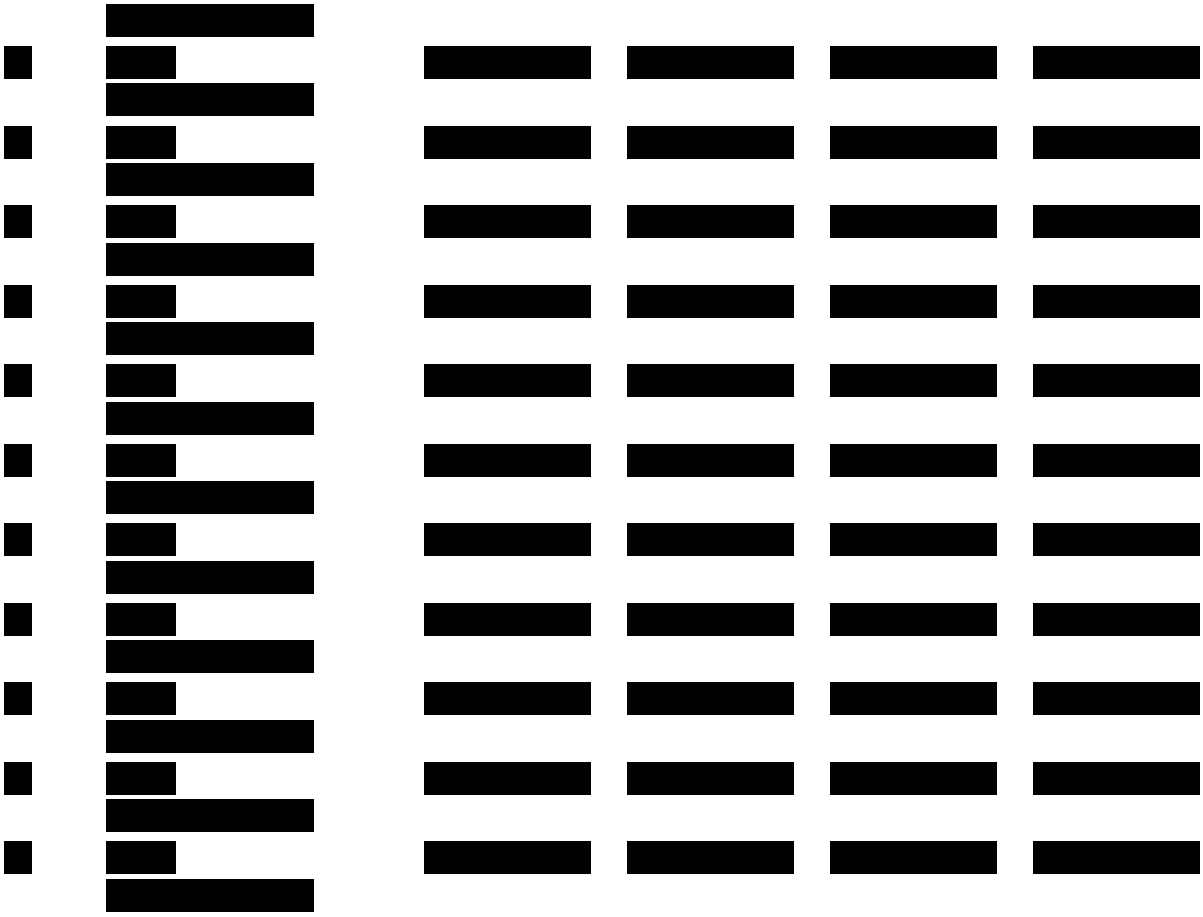


Table 81 Statistical fits – DIC fractional polynomial fixed-effect

FP	parameters	OS	PFS
1st	P0	5751.9	9567.8
1st	P1	5855.9	9579.3
1st	P05	5775.4	9586.0
1st	Pm1	5661.9	9355.8
1st	Pm05	5708.3	9483.4
2nd	Pm050	5611.8	NA
2nd	Pm10	5599.3	8913.2
2nd	Pm1m1	5596.7	8837.8
2nd	Pm105	5605.8	8975.5
2nd	Pm11	5614.4	9044.3

Fractional polynomial models based on random effect

OS fractional polynomial 1st order

Figure 143 OS fractional polynomial model survival curve – random-effect model (P=0)

Figure 144 OS fractional polynomial model HR plot – random-effect model (P=0)

Figure 145 OS fractional polynomial model survival curve – random-effect model (P=1)

Figure 146 OS fractional polynomial model HR plot – random-effect model (P=1)

Figure 147 OS fractional polynomial model survival curve – random-effect model (P=0.5)

Figure 148 OS fractional polynomial model HR plot – random-effect model (P=0.5)

Figure 149 OS fractional polynomial model survival curve – random-effect model (P=-1)

Figure 150 OS fractional polynomial model HR plot – random-effect model (P=-1)

Figure 151 OS fractional polynomial model survival curve – random-effect model (P=-0.5)

Figure 152 OS fractional polynomial model HR plot – random-effect model (P=-0.5)

Table 82 OS fractional polynomial model HR plot – random-effect model BEV+IFN

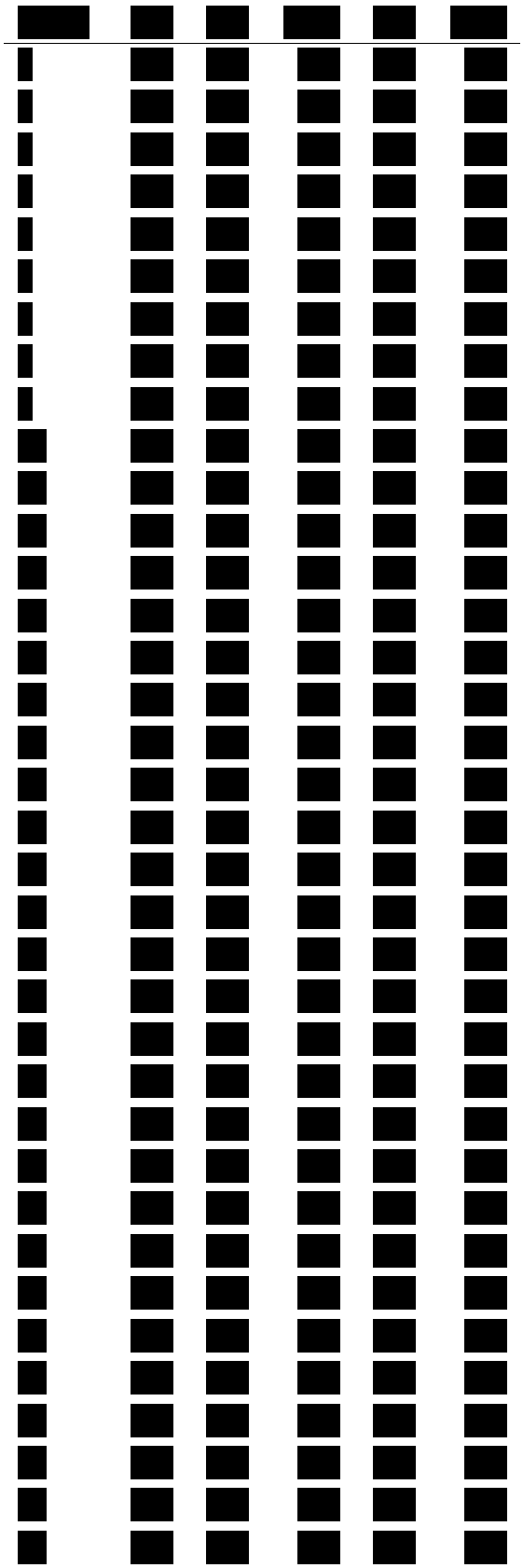
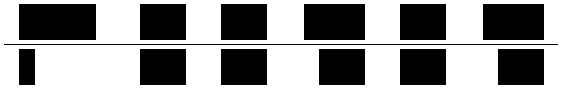


Table 83 OS fractional polynomial model HR plot – random-effect model IFN



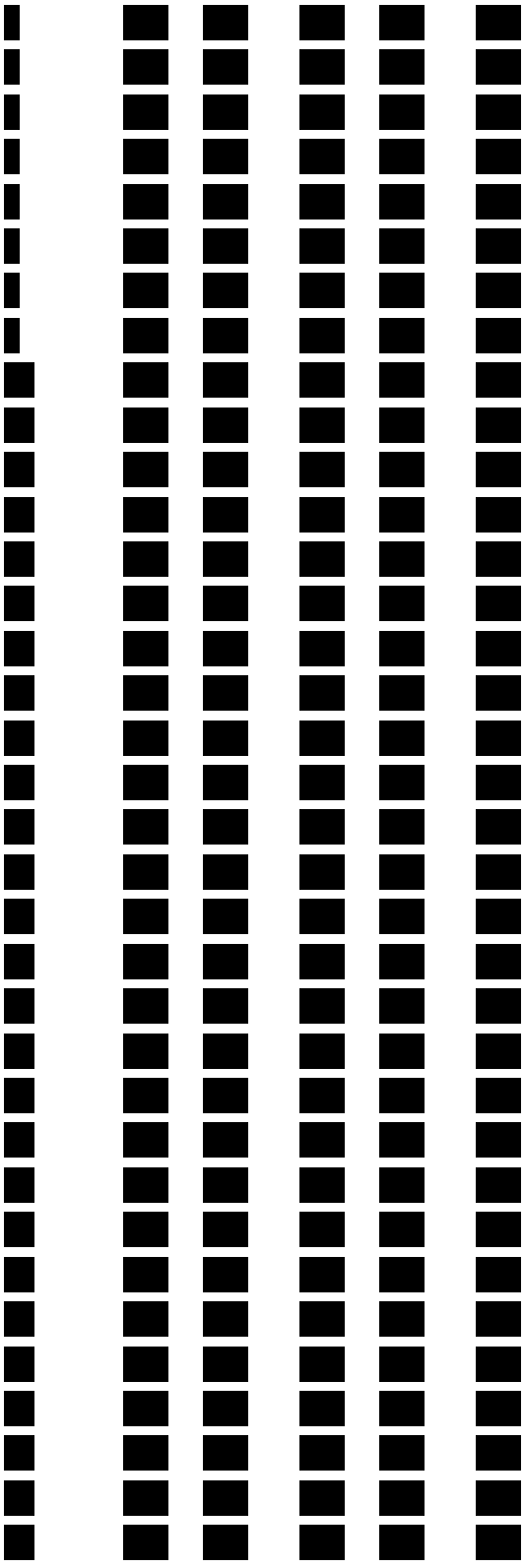
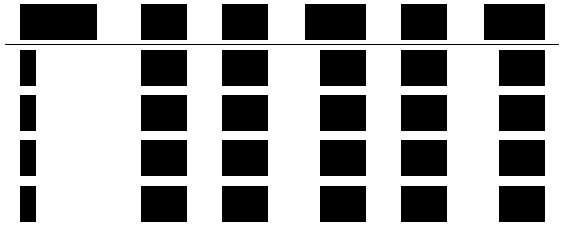


Table 84 OS fractional polynomial model HR plot – random-effect model pazopanib



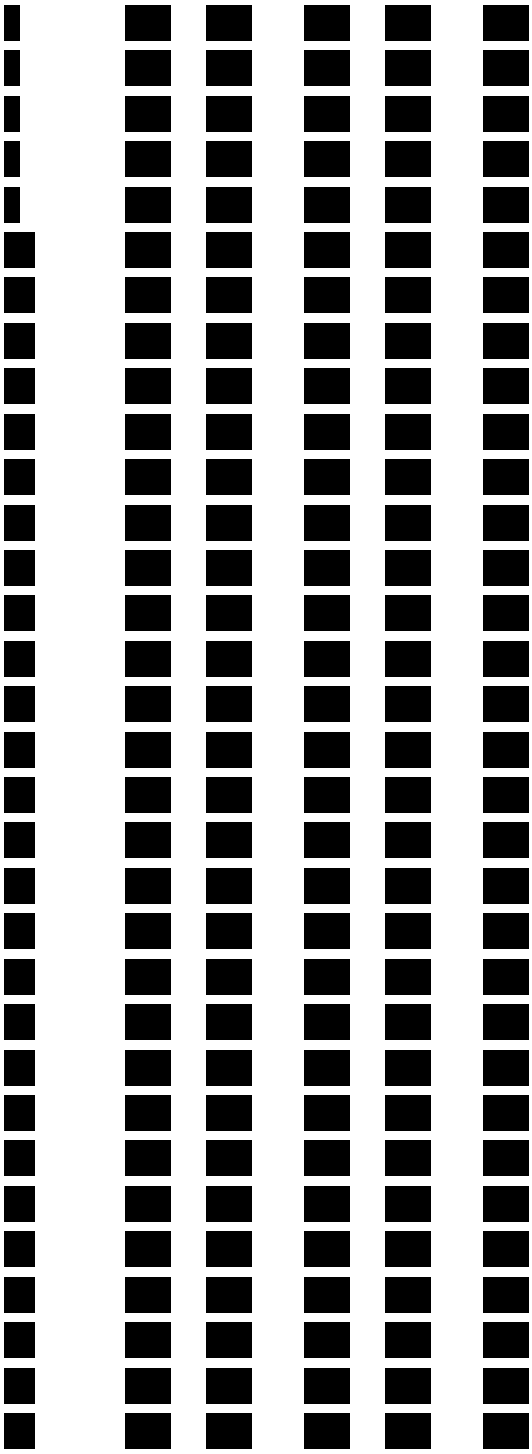
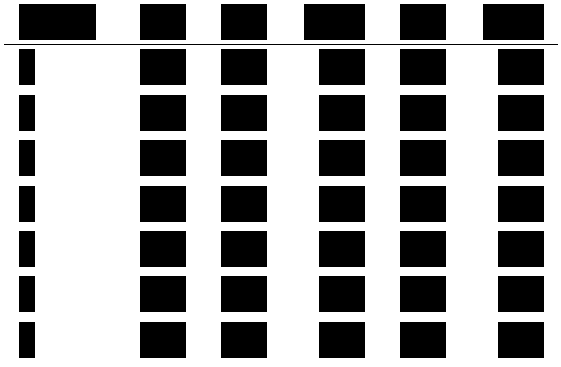


Table 85 OS fractional polynomial model HR plot – random-effect model sorafenib



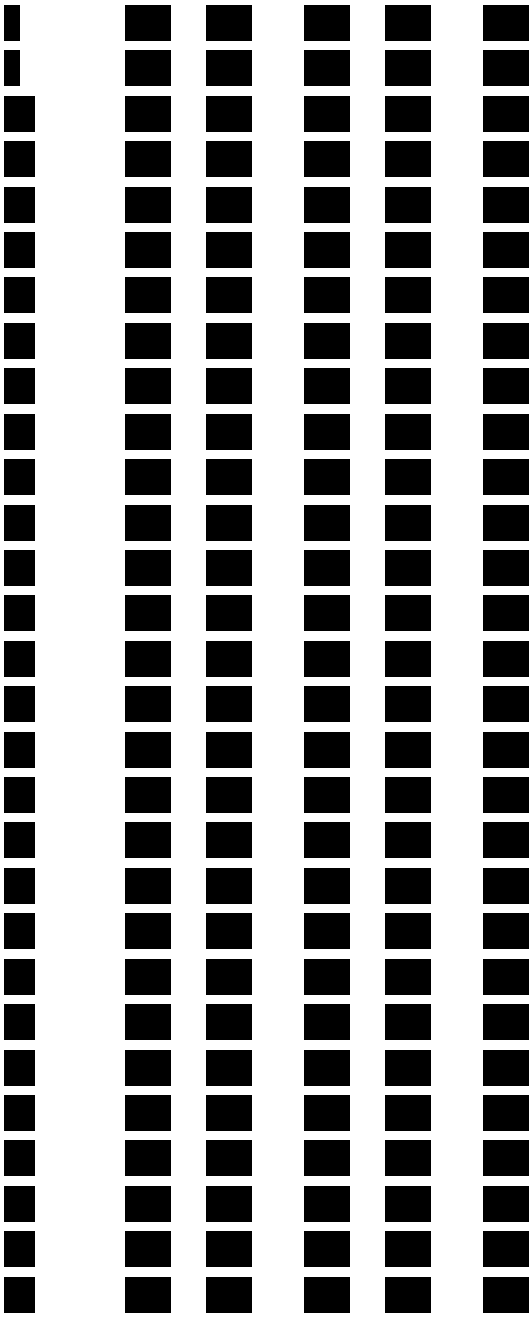
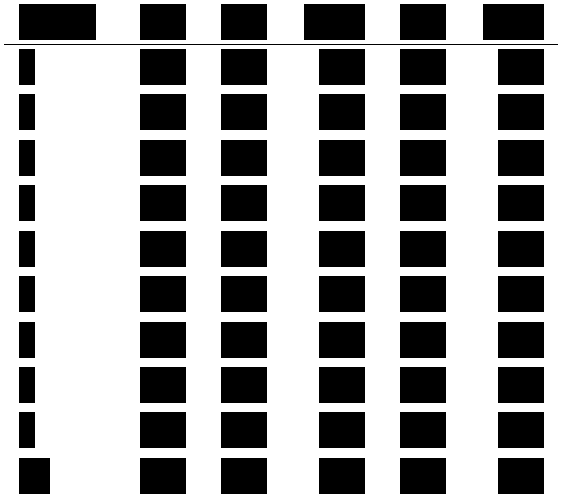


Table 86 OS fractional polynomial model HR plot – random-effect model sunitinib



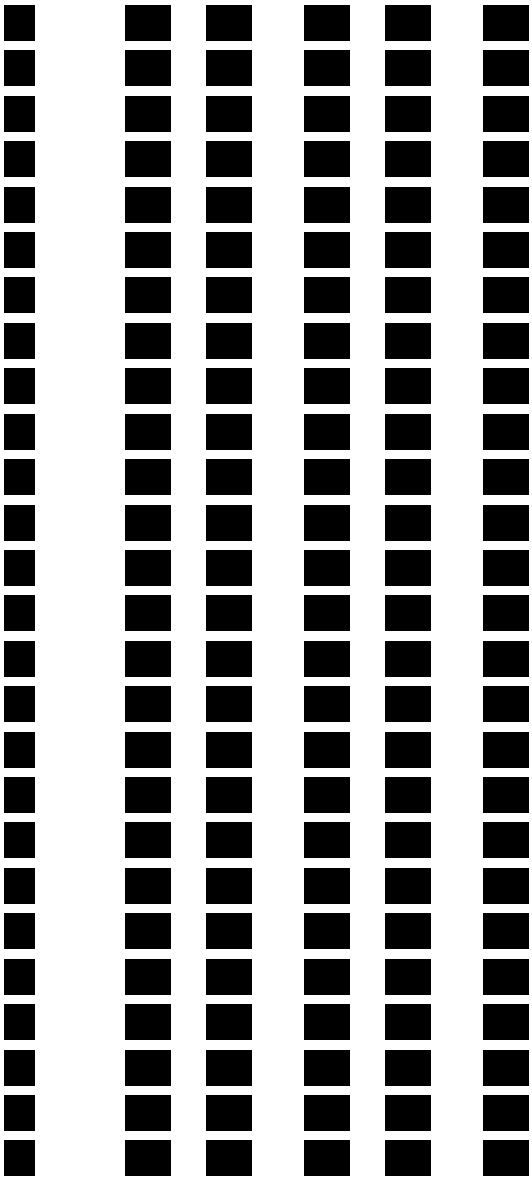
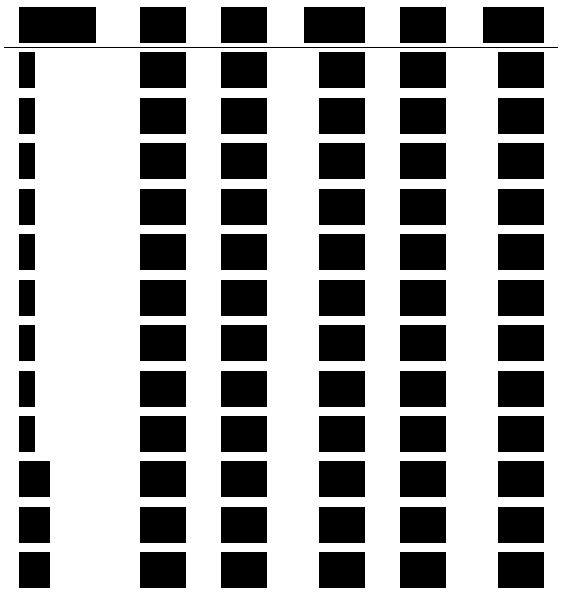


Table 87 OS fractional polynomial model HR plot – random-effect model temsirolimus



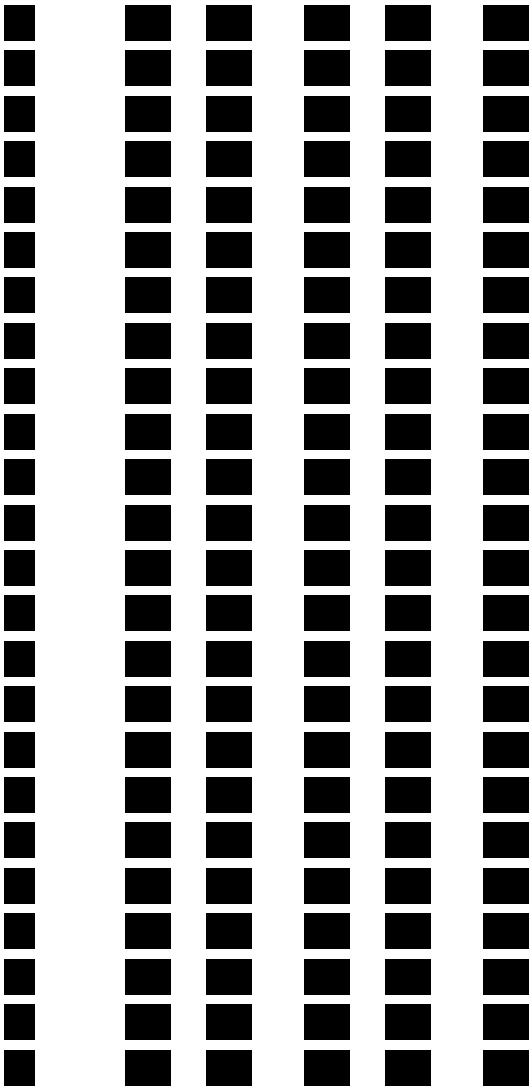
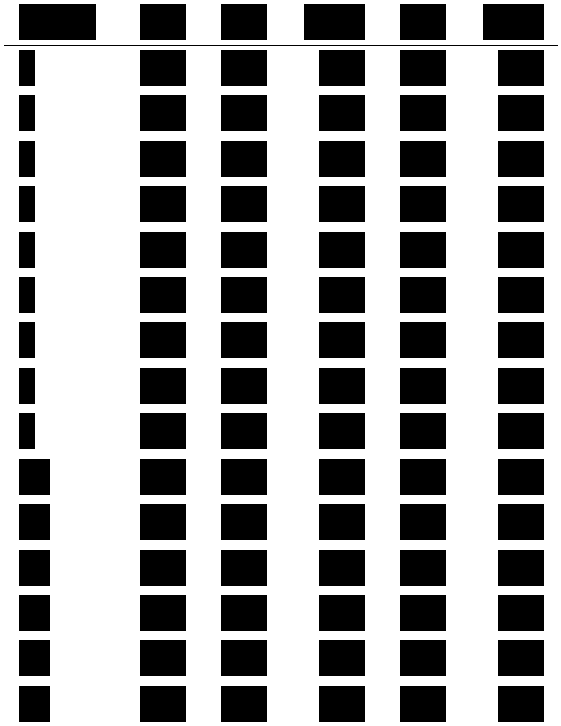
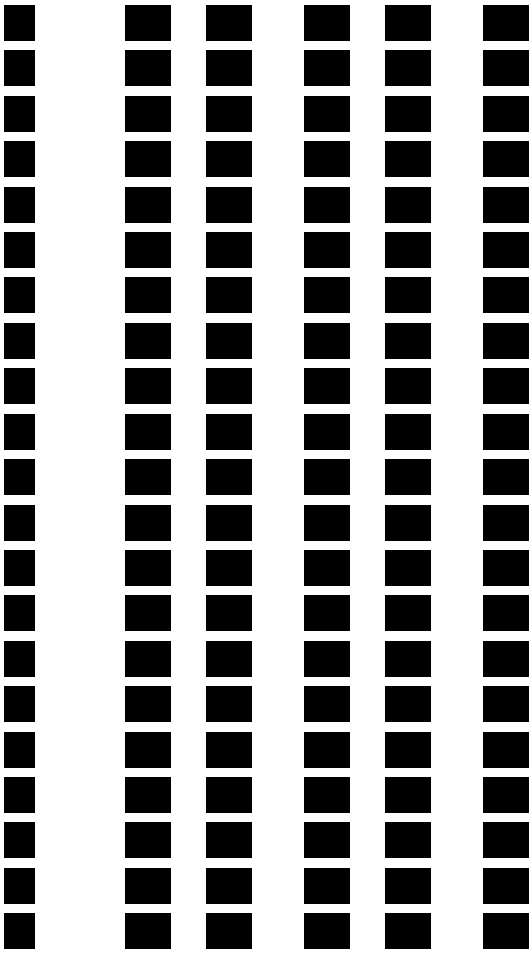


Table 88 OS fractional polynomial model HR plot – random-effect model tivozanib





OS fractional polynomial 2nd order

Figure 153 OS fractional polynomial model survival curve – random-effect model (P1=-0.5, P2=0)

Figure 154 OS fractional polynomial model HR plot – random-effect model (P1=-0.5, P2=0)

Figure 155 OS fractional polynomial model survival curve – random-effect model (P1=-1, P2=0)

Figure 156 OS fractional polynomial model HR plot – random-effect model (P1=-1, P2=0)

Figure 157 OS fractional polynomial model survival curve – random-effect model (P1=-1, P2=-1)

Figure 158 OS fractional polynomial model HR plot – random-effect model (P1=-1, P2=-1)

Figure 159 OS fractional polynomial model survival curve – random-effect model (P1=-1, P2=0.5)

Figure 160 OS fractional polynomial model HR plot – random-effect model (P1=-1, P2=0.5)

Figure 161 OS fractional polynomial model survival curve – random-effect model (P1=-1, P2=1)

Figure 162 OS fractional polynomial model HR plot – random-effect model (P1=-1, P2=1)

Table 89 OS fractional polynomial model HR plot – random-effect model BEV+IFN

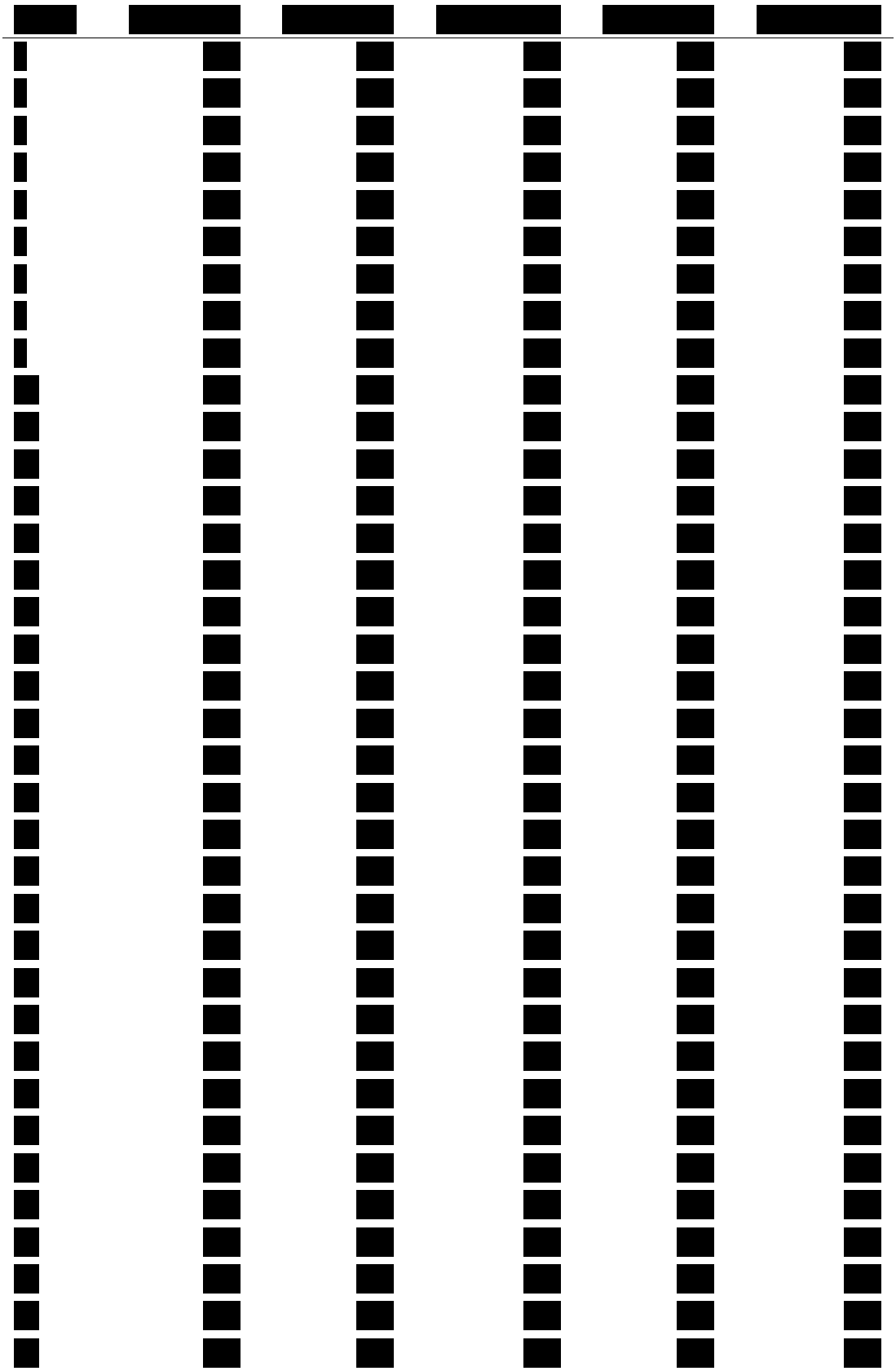


Table 90 OS fractional polynomial model HR plot – random-effect model IFN

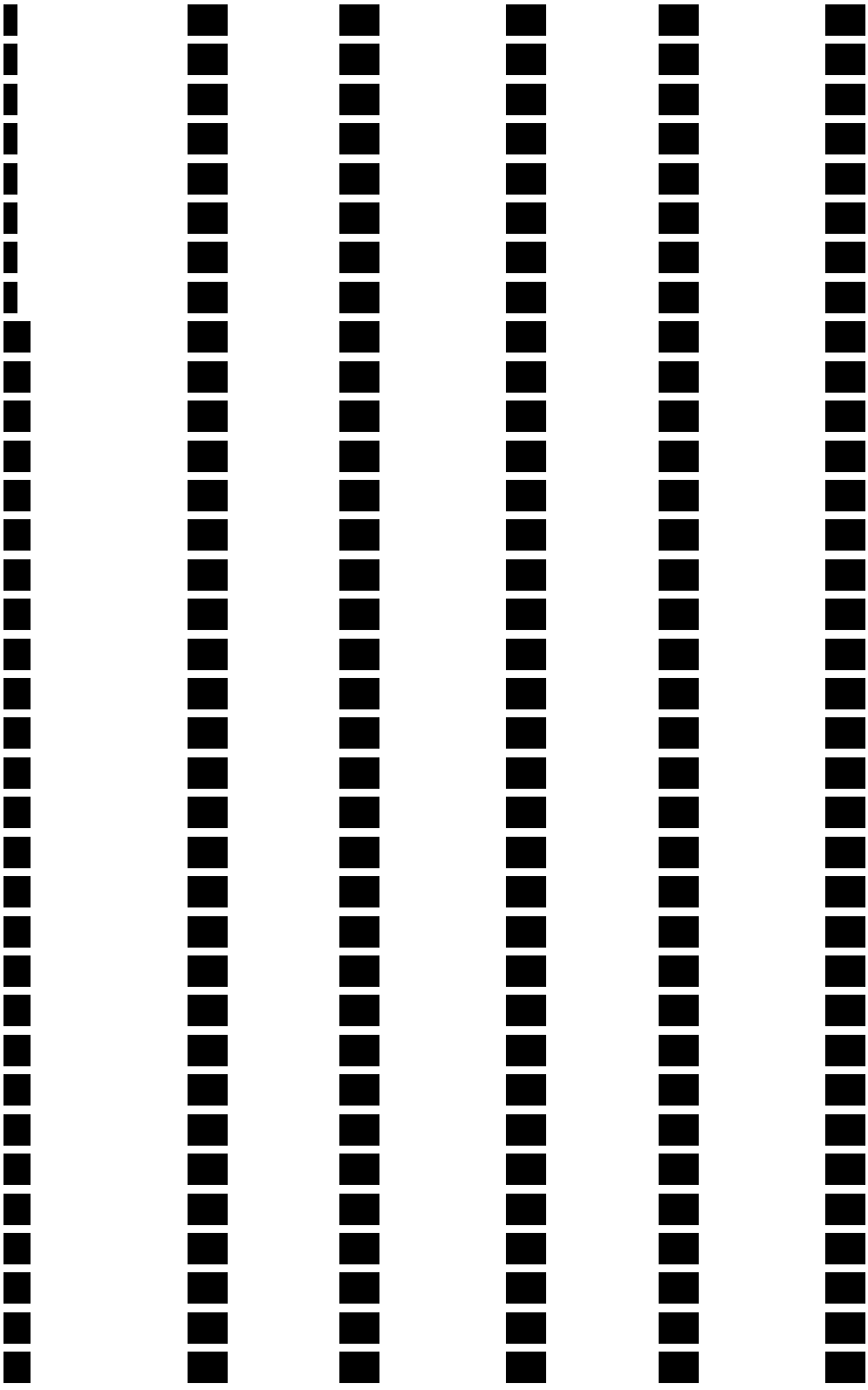
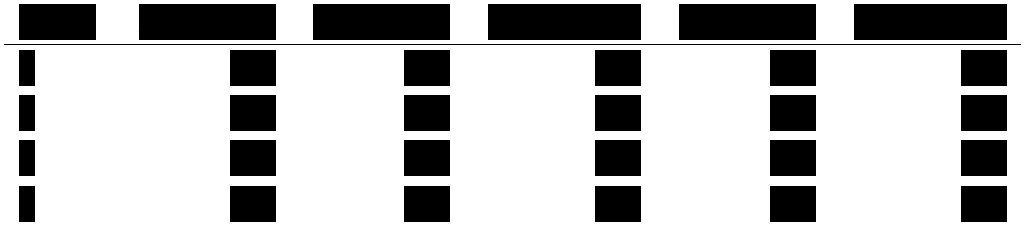


Table 91 OS fractional polynomial model HR plot – random-effect model pazopanib



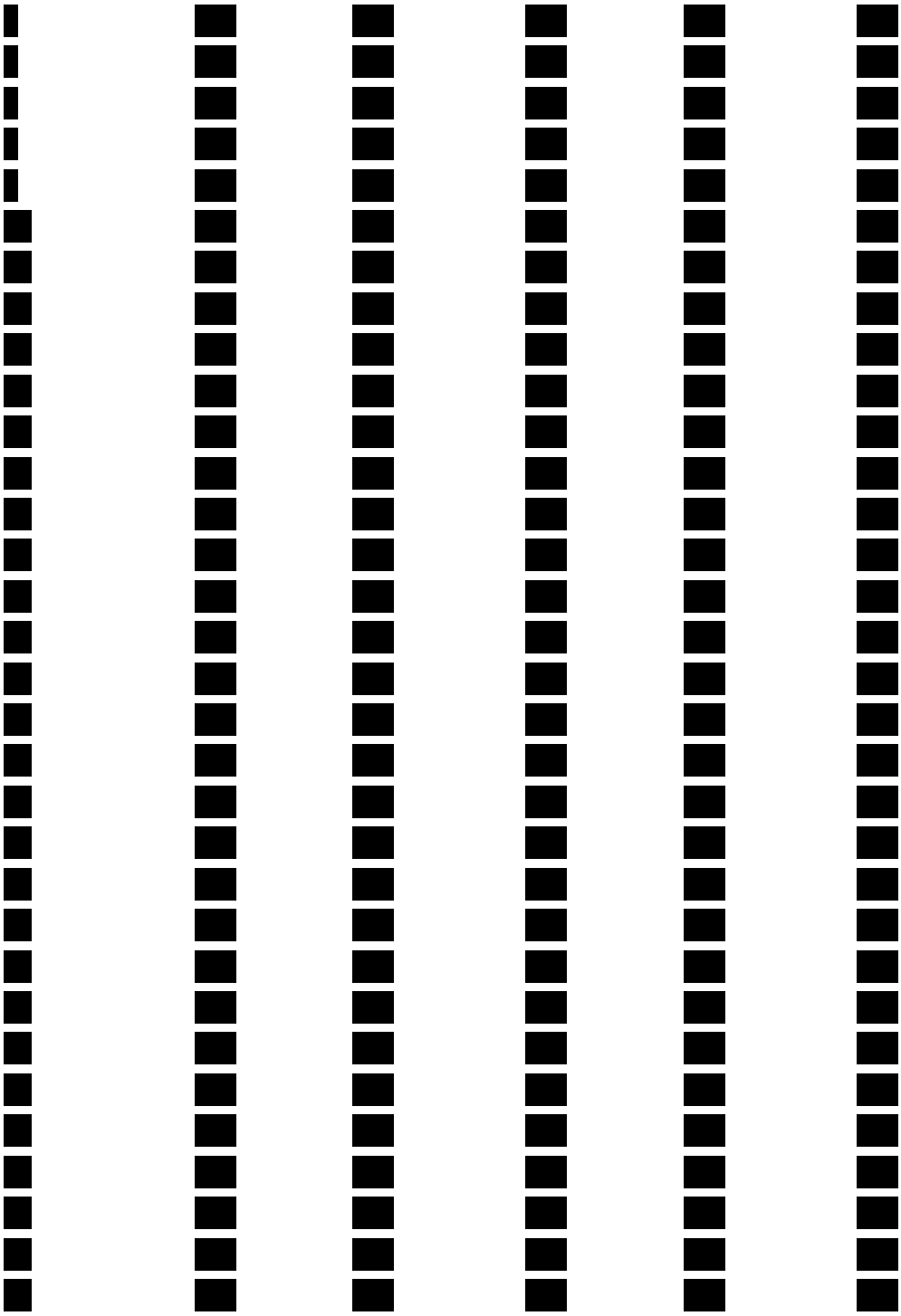
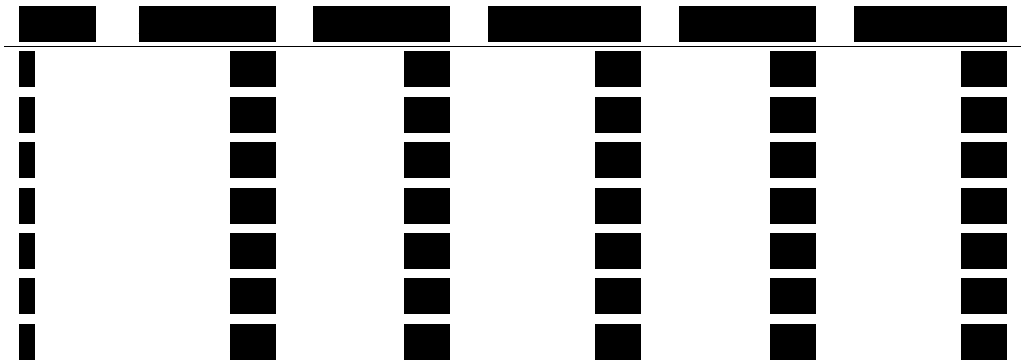


Table 92 OS fractional polynomial model HR plot – random-effect model sorafenib



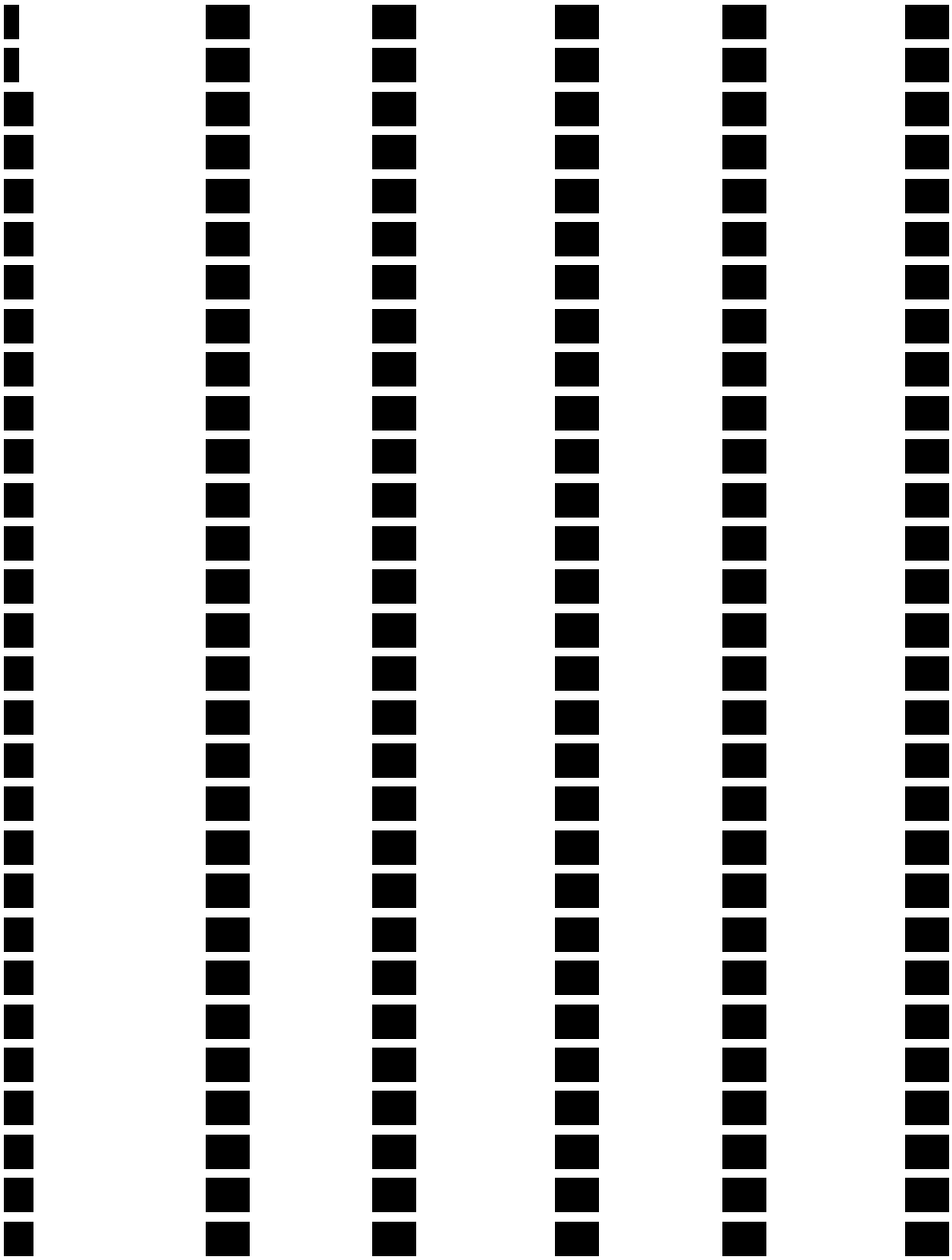
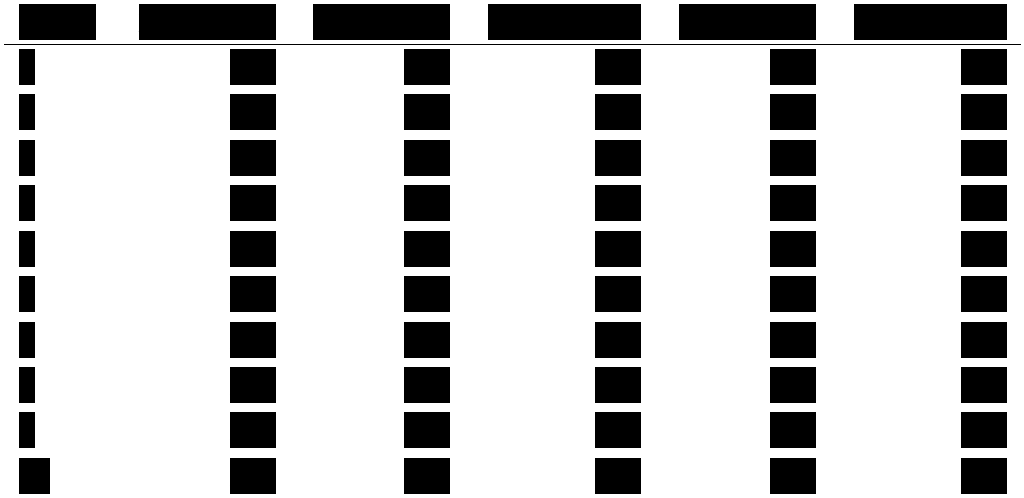


Table 93 OS fractional polynomial model HR plot – random-effect model sunitinib



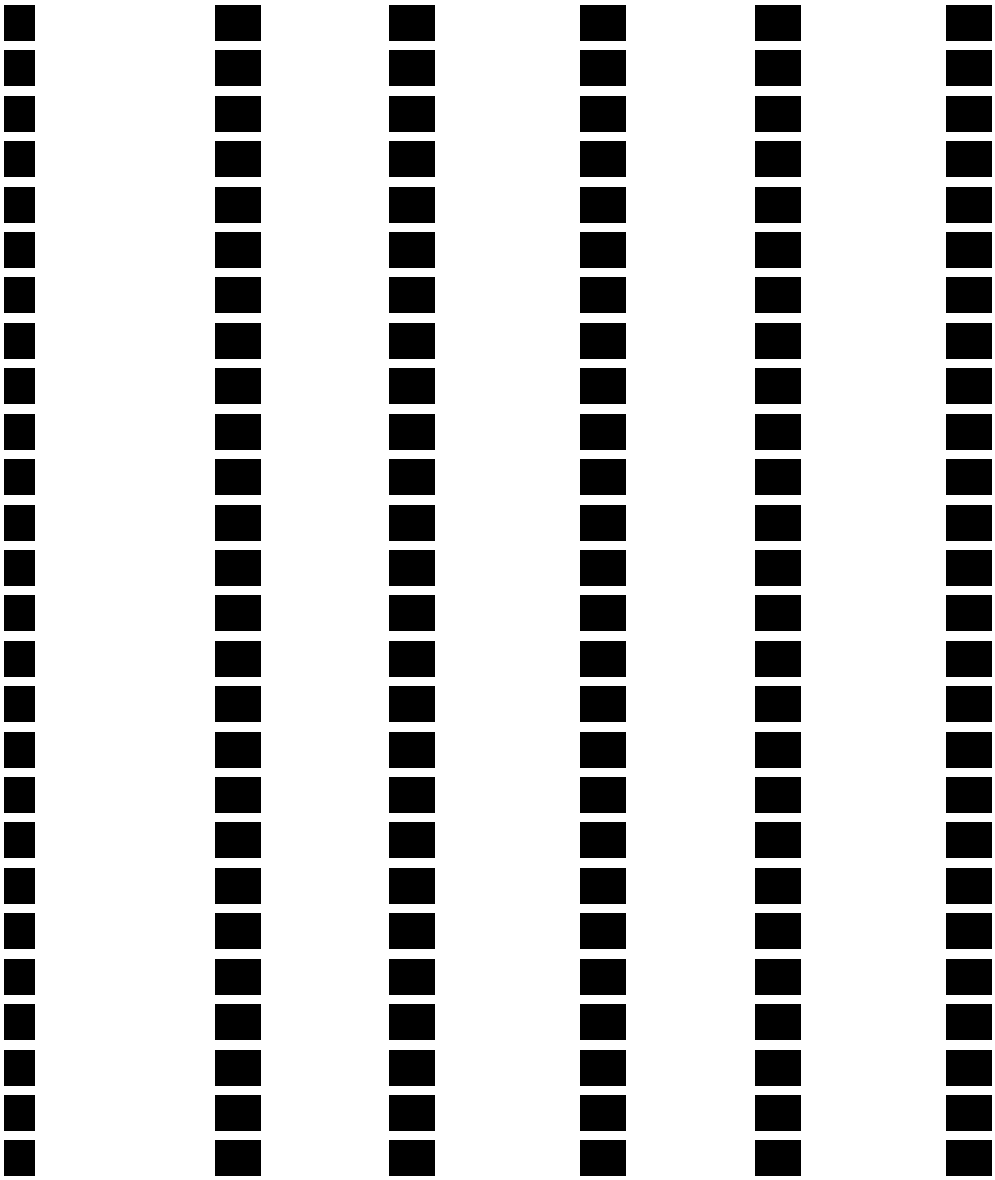
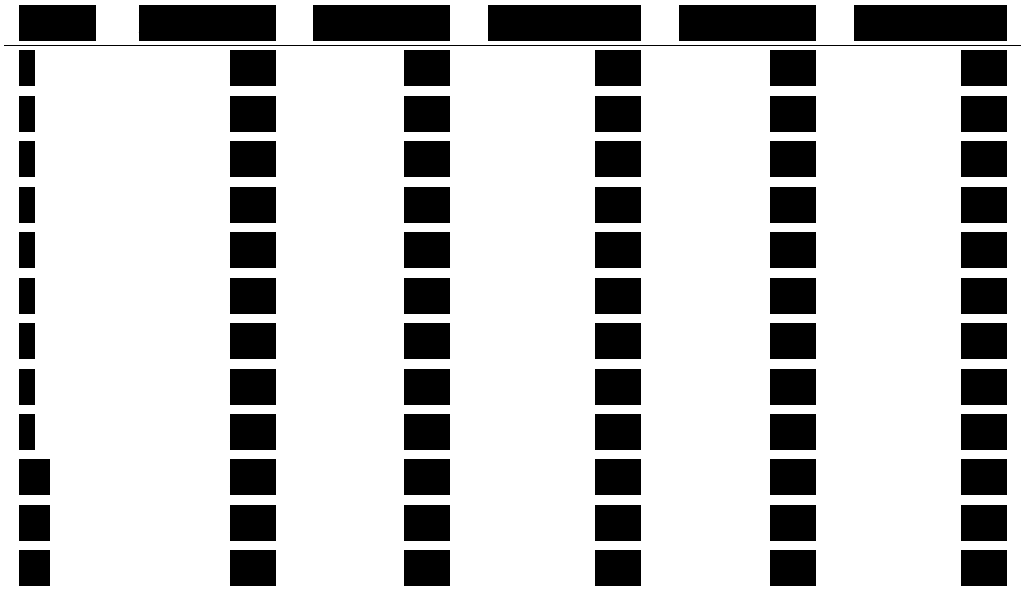


Table 94 OS fractional polynomial model HR plot – random-effect model temsirolimus



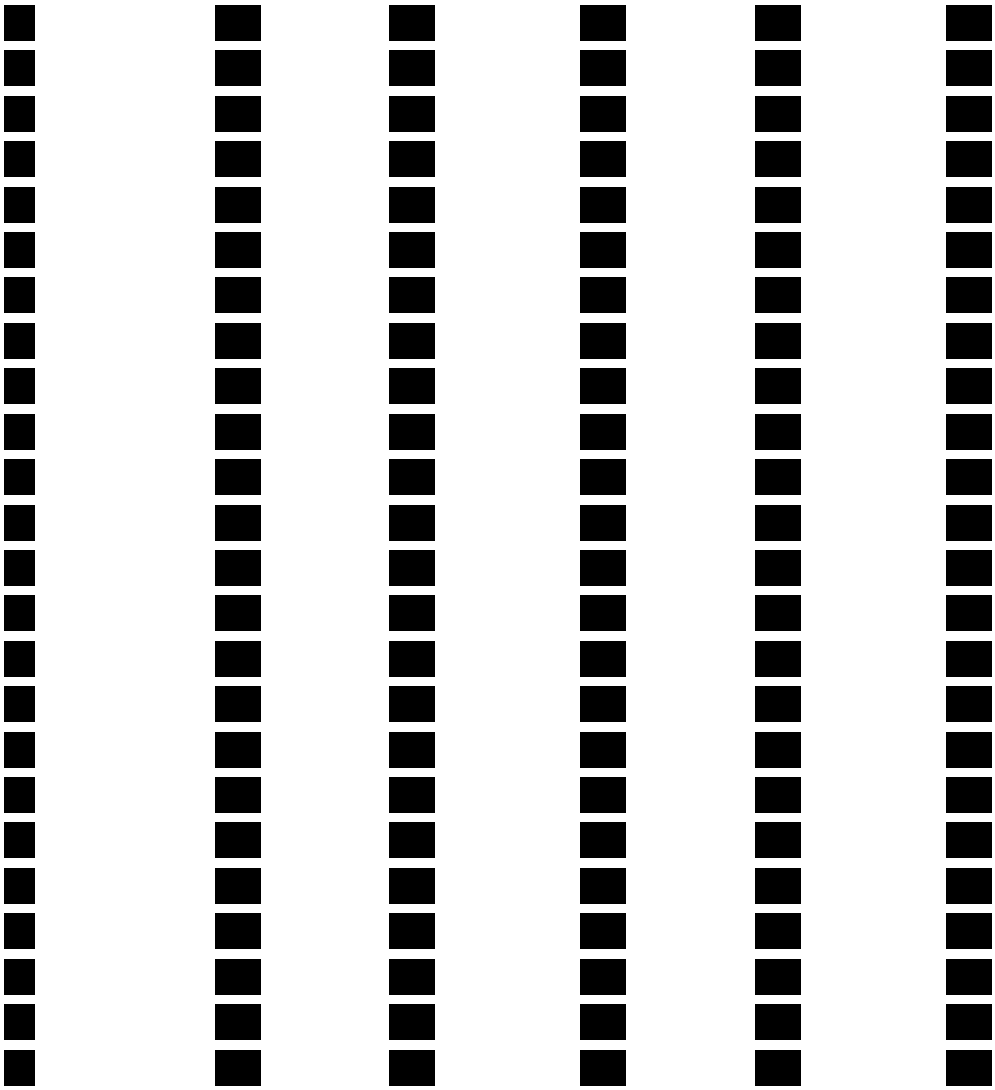
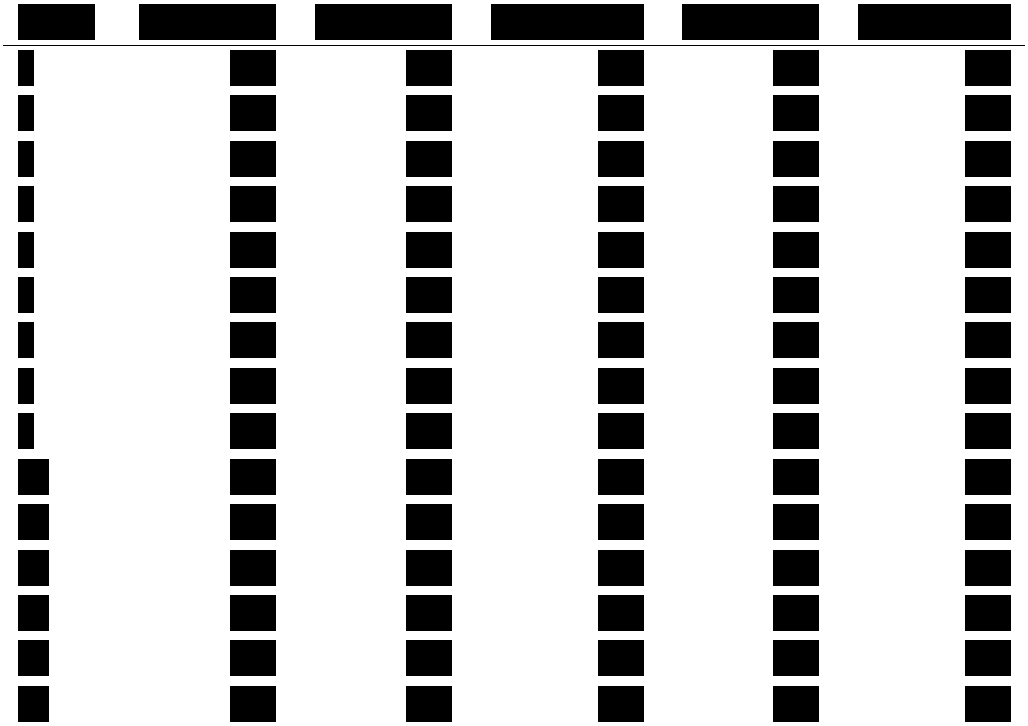
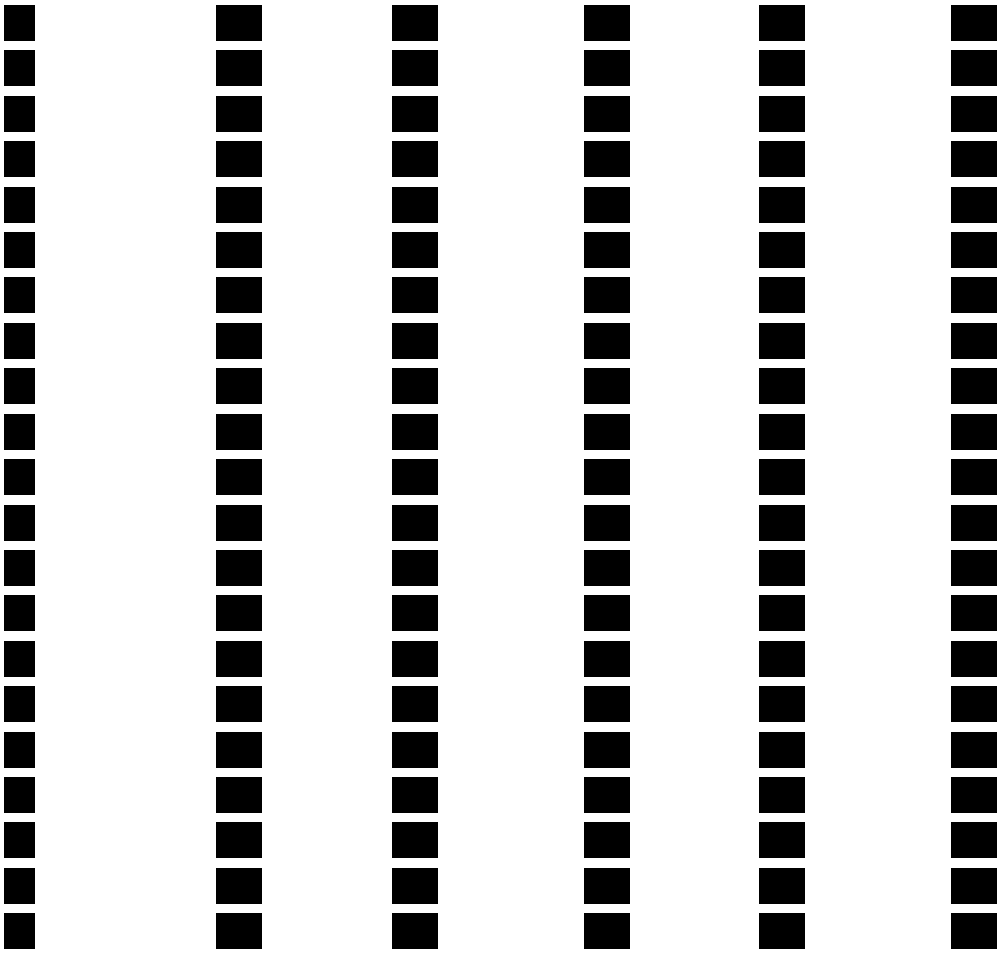


Table 95 OS fractional polynomial model HR plot – random-effect model tivozanib





PFS fractional polynomial 1st order

Figure 163 PFS fractional polynomial model survival curve – random-effect model (P=0)

Figure 164 PFS fractional polynomial model HR plot – random-effect model (P=0)

Figure 165 PFS fractional polynomial model survival curve – random-effect model (P=1)

Figure 166 PFS fractional polynomial model HR plot – random-effect model (P=1)

Figure 167 PFS fractional polynomial model survival curve – random-effect model (P=0.5)

Figure 168 PFS fractional polynomial model HR plot – random-effect model (P=0.5)

Figure 169 PFS fractional polynomial model survival curve – random-effect model (P=-1)

Figure 170 PFS fractional polynomial model HR plot – random-effect model (P=-1)

Figure 171 PFS fractional polynomial model survival curve – random-effect model (P=-0.5)

Figure 172 PFS fractional polynomial model HR plot – random-effect model (P=-0.5)

Table 96 PFS fractional polynomial model HR plot – random-effect model BEV+IFN

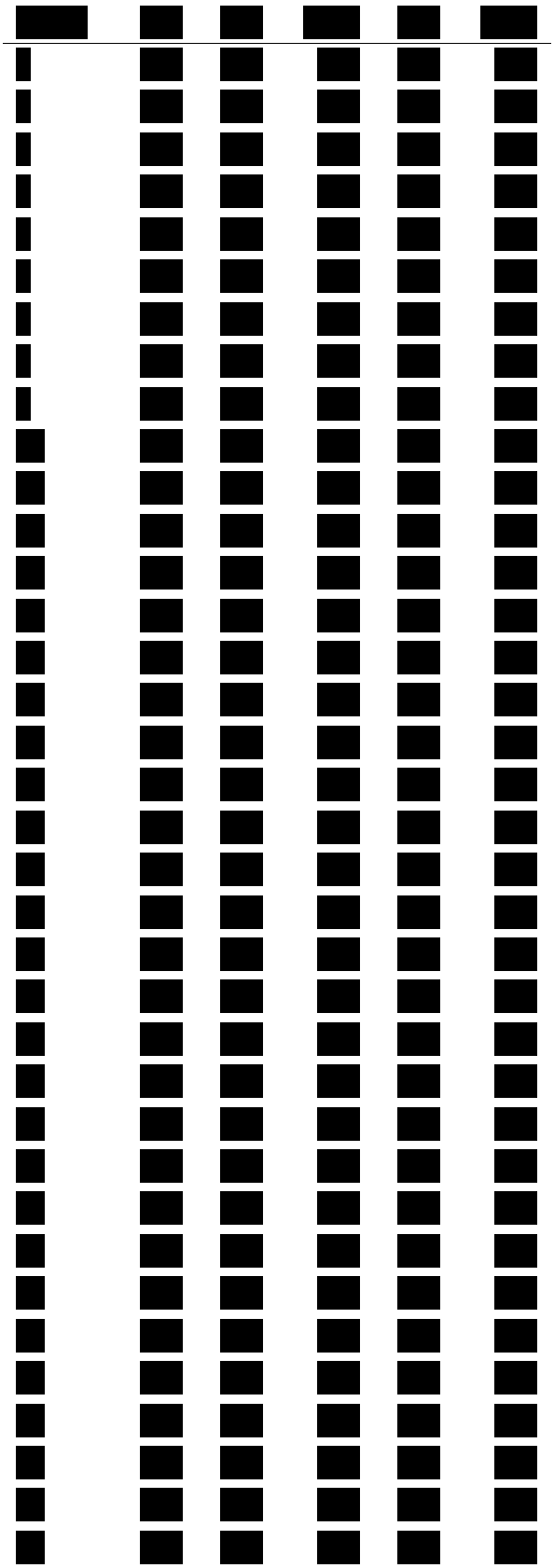
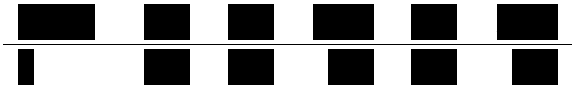


Table 97 PFS fractional polynomial model HR plot – random-effect model IFN



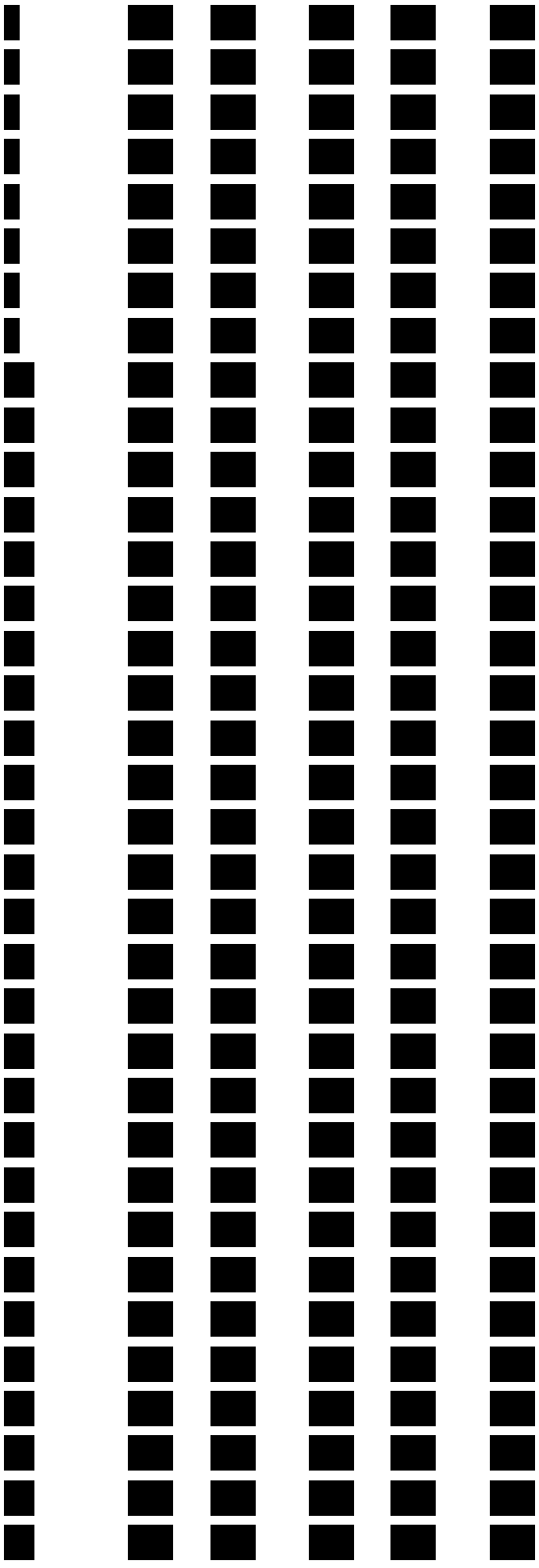
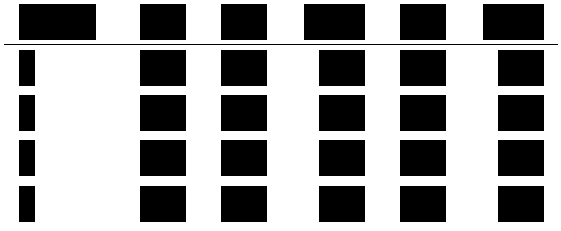


Table 98 PFS fractional polynomial model HR plot – random-effect model pazopanib



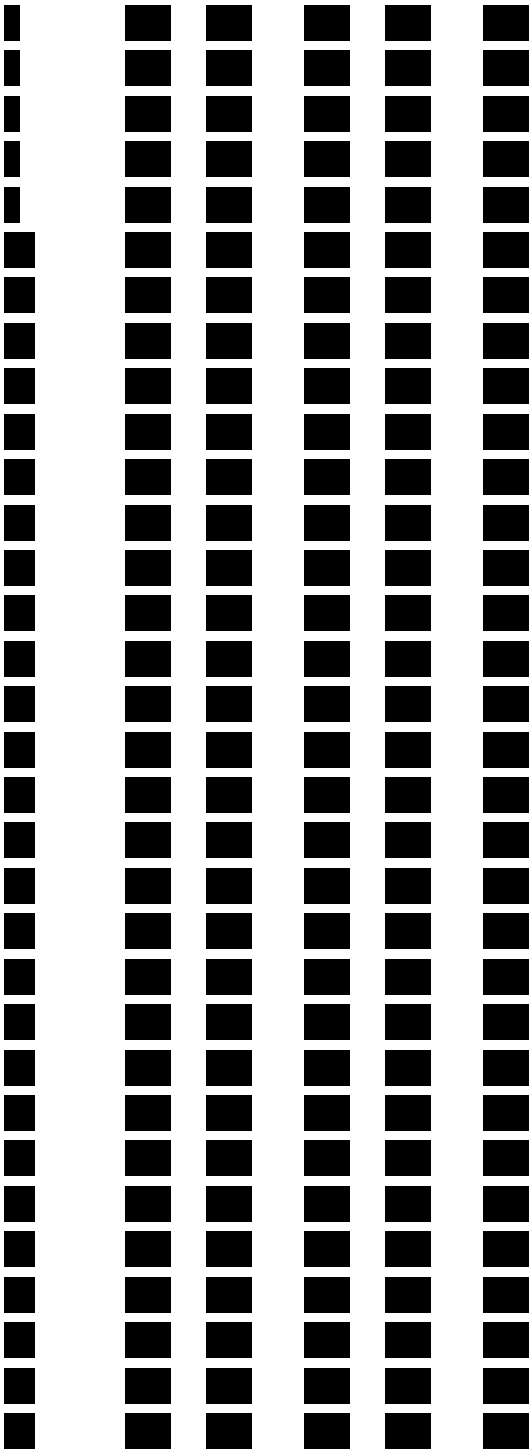
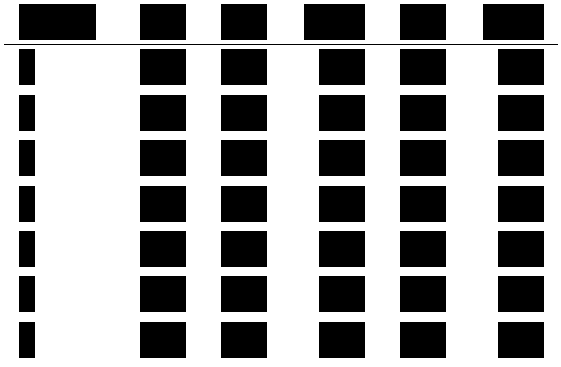


Table 99 PFS fractional polynomial model HR plot – random-effect model placebo



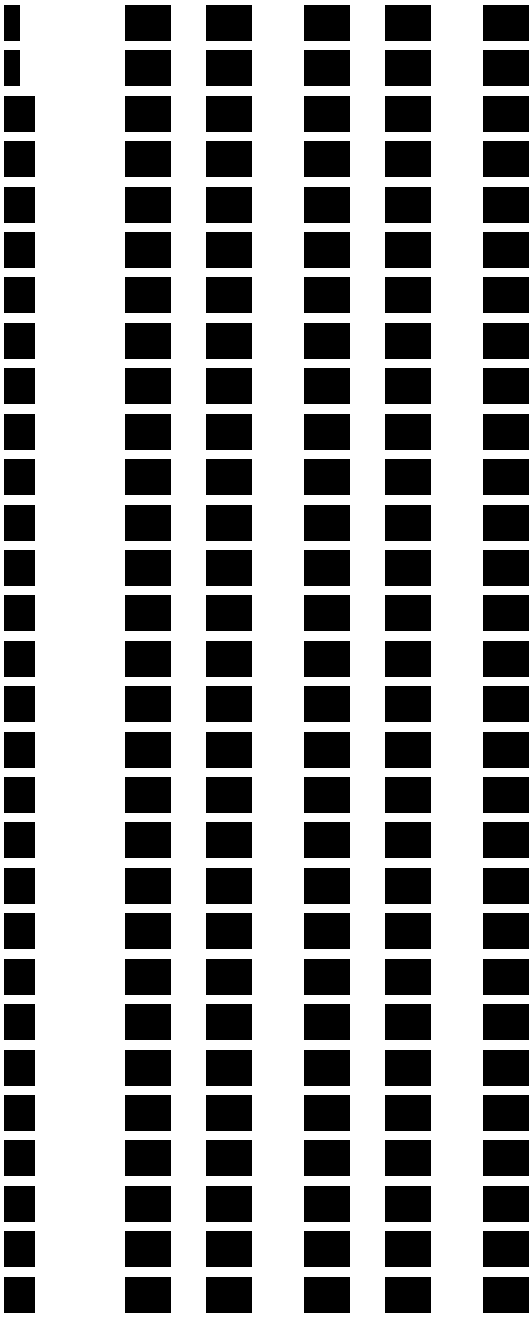
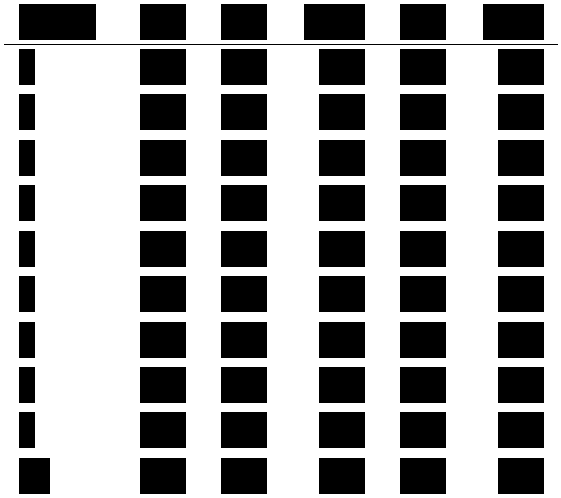


Table 100 PFS fractional polynomial model HR plot – random-effect model sorafenib



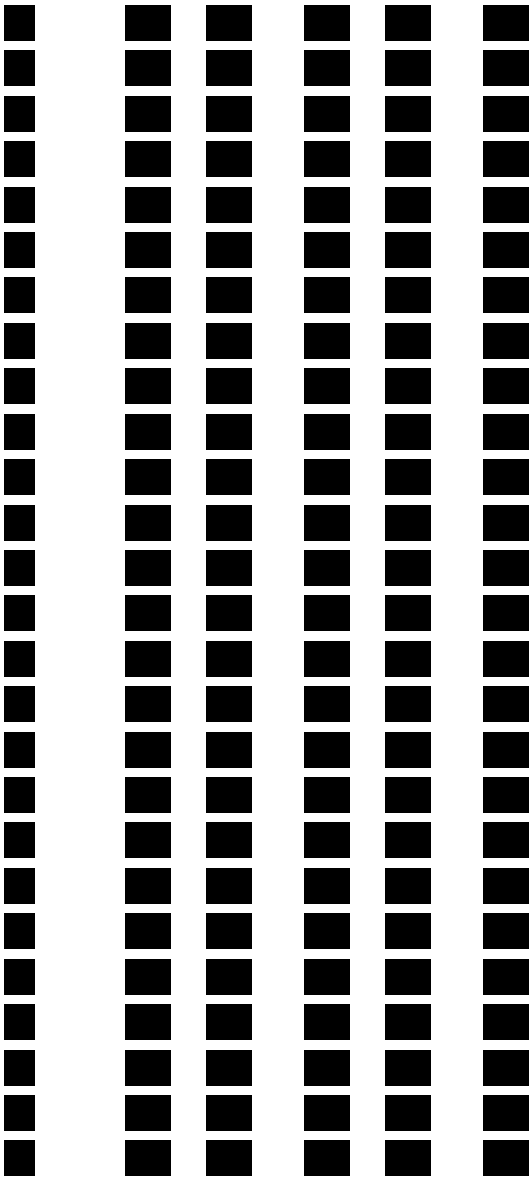
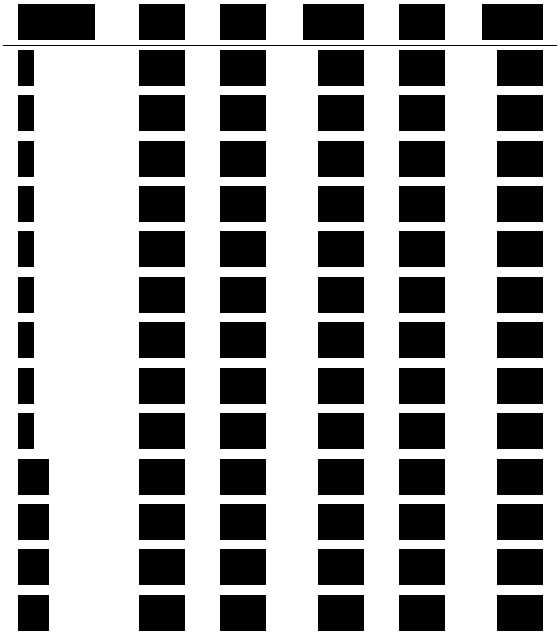


Table 101 PFS fractional polynomial model HR plot – random-effect model sunitinib



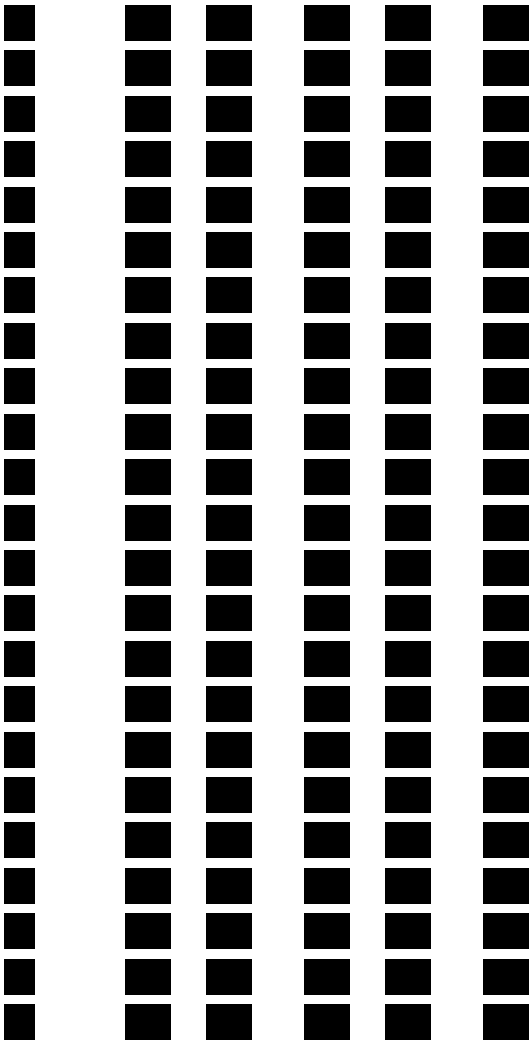
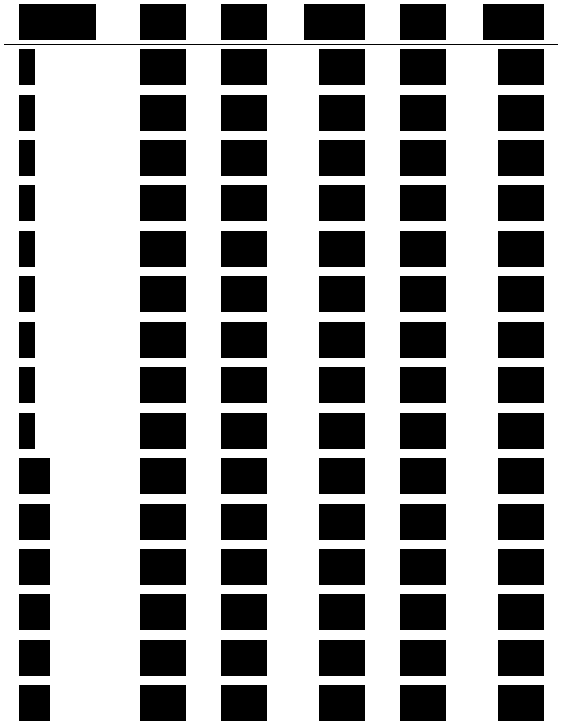


Table 102 PFS fractional polynomial model HR plot – random-effect model temsirolimus



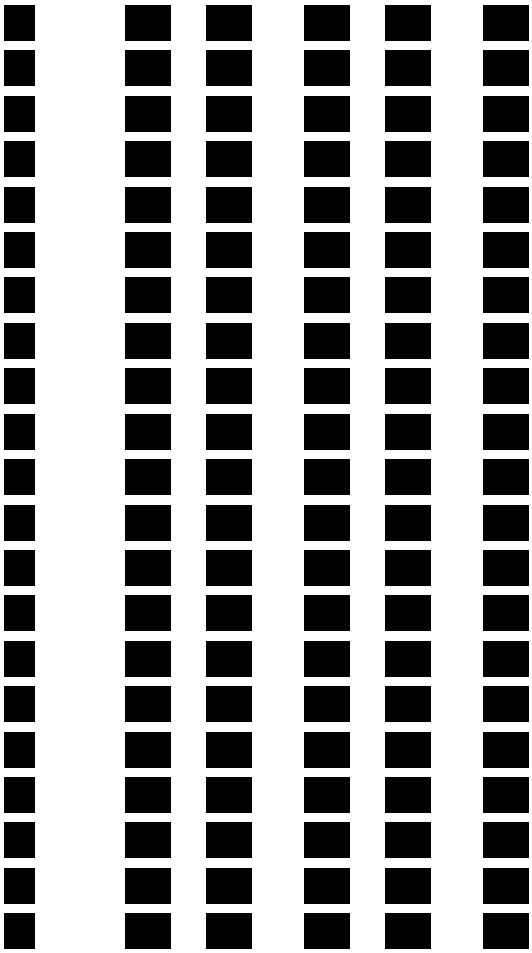
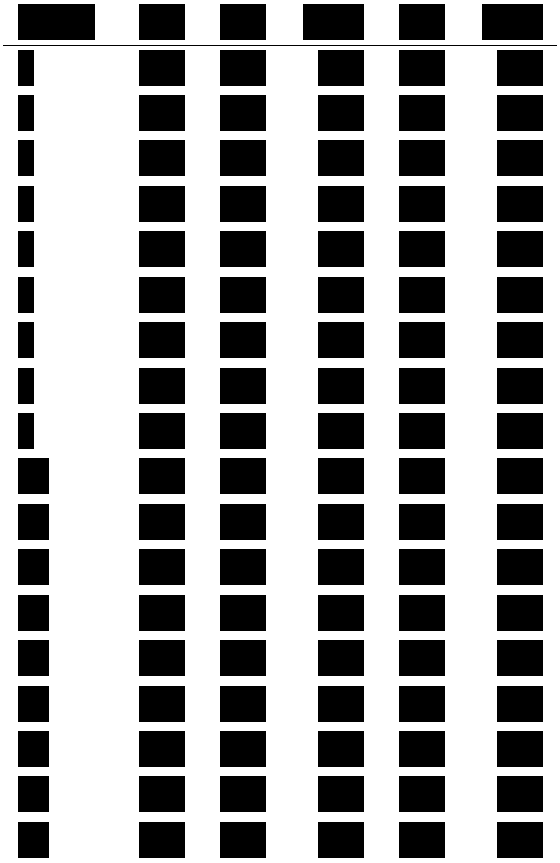


Table 103 PFS fractional polynomial model HR plot – random-effect model tivozanib



PFS fractional polynomial 2nd order

Figure 173 PFS fractional polynomial model survival curve – random-effect model (P1=-0.5, P2=0)

Figure 174 PFS fractional polynomial model HR plot – random-effect model (P1=-0.5, P2=0)

Figure 175 PFS fractional polynomial model survival curve – random-effect model (P1=-1, P2=0)

Figure 176 PFS fractional polynomial model HR plot – random-effect model (P1=-1, P2=0)

Figure 177 PFS fractional polynomial model survival curve – random-effect model (P1=-1, P2=-1)

Figure 178 PFS fractional polynomial model HR plot – random-effect model (P1=-1, P2=-1)

Figure 179 PFS fractional polynomial model survival curve – random-effect model (P1=-1, P2=0.5)

Figure 180 PFS fractional polynomial model HR plot – random-effect model (P1=-1, P2=0.5)

Figure 181 PFS fractional polynomial model survival curve – random-effect model (P1=-1, P2=1)

Figure 182 PFS fractional polynomial model HR plot – random-effect model (P1=-1, P2=1)

Table 104 PFS fractional polynomial model HR plot – random-effect model BEV+IFN

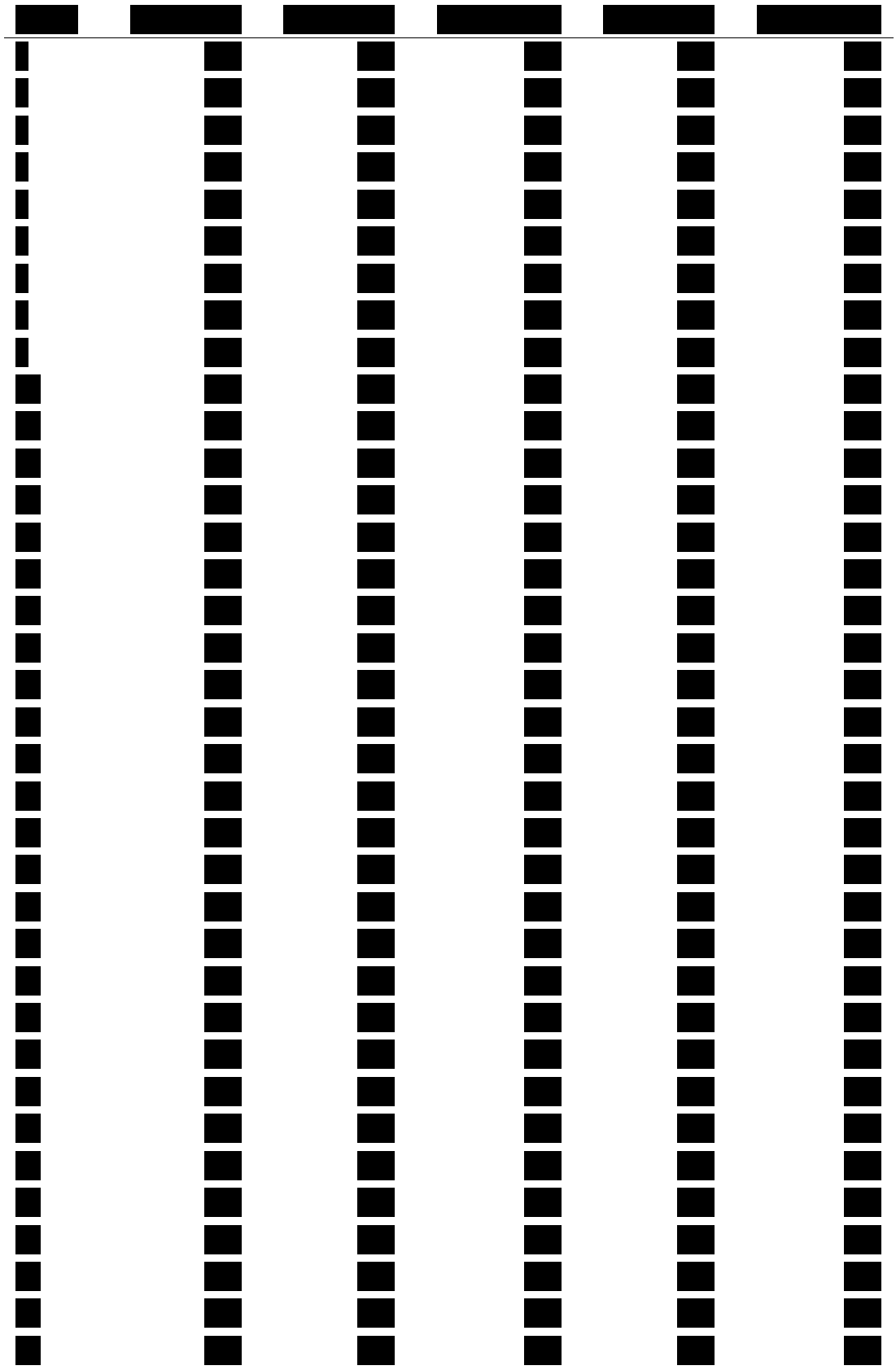


Table 105 PFS fractional polynomial model HR plot – random-effect model IFN

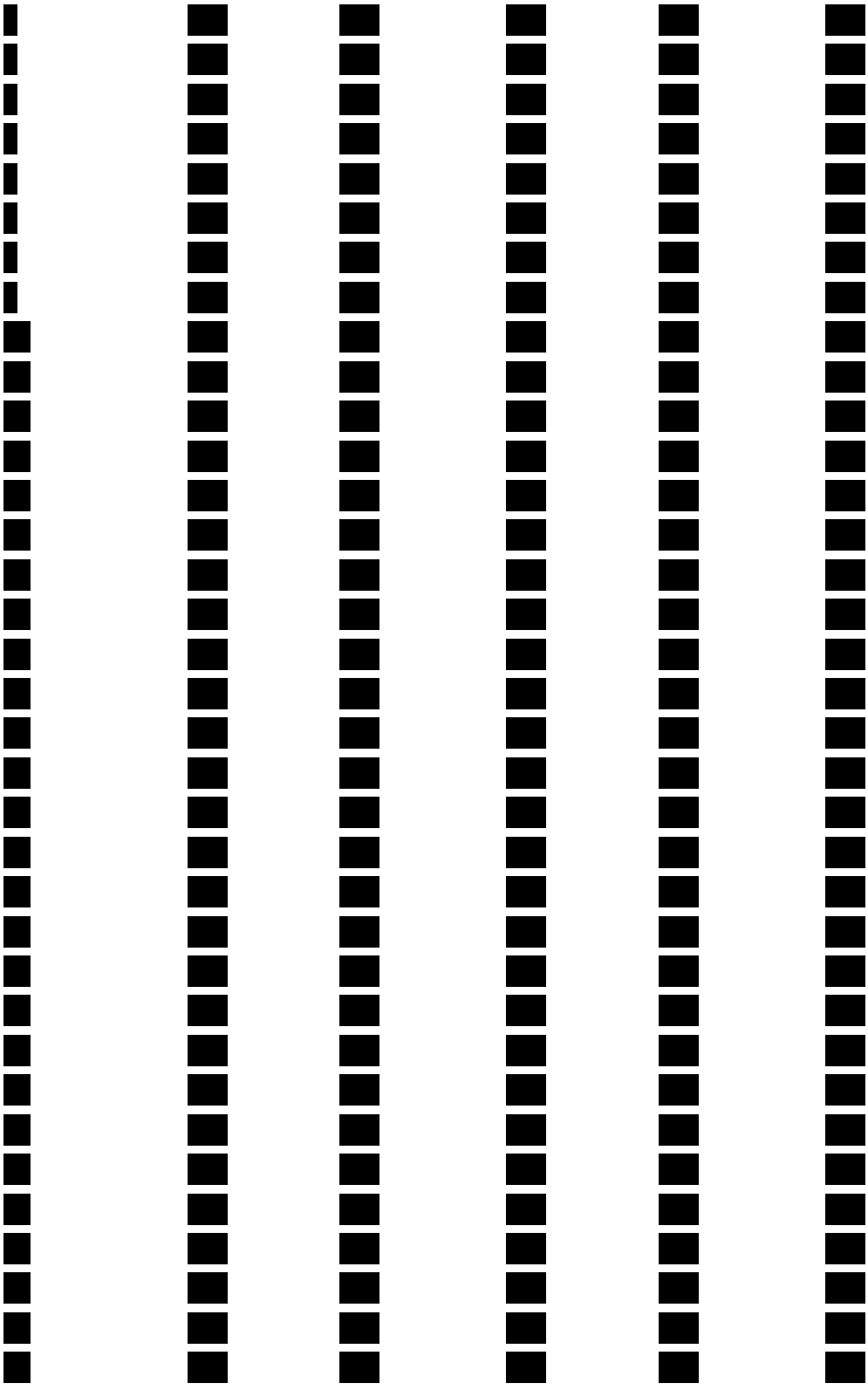
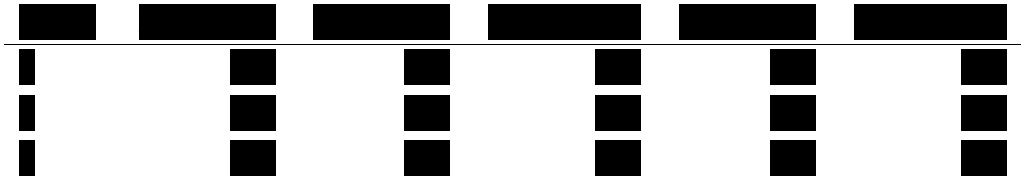


Table 106 PFS fractional polynomial model HR plot – random-effect model pazopanib



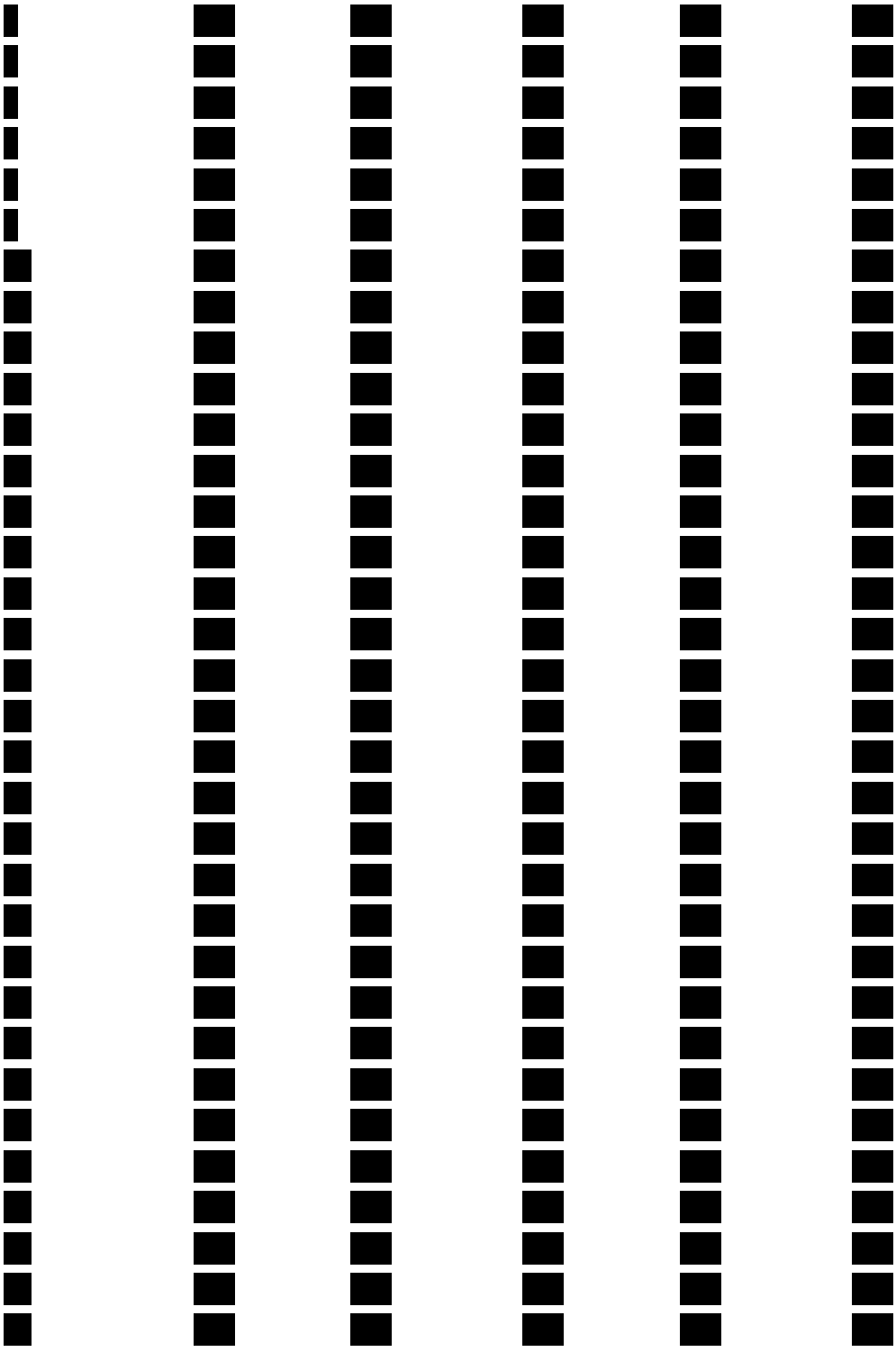
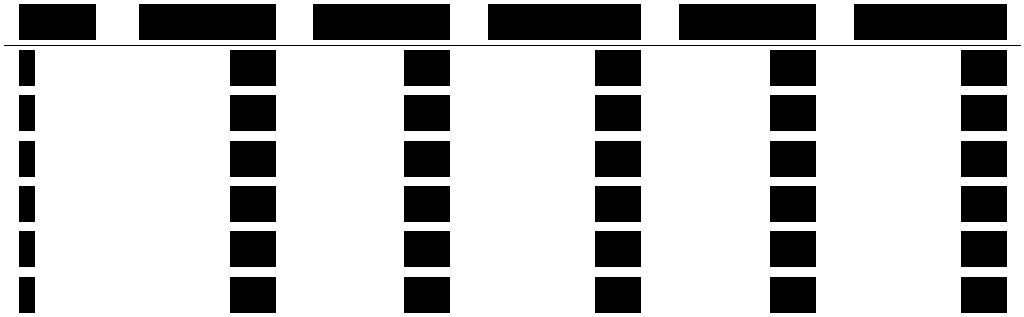


Table 107 PFS fractional polynomial model HR plot – random-effect model placebo



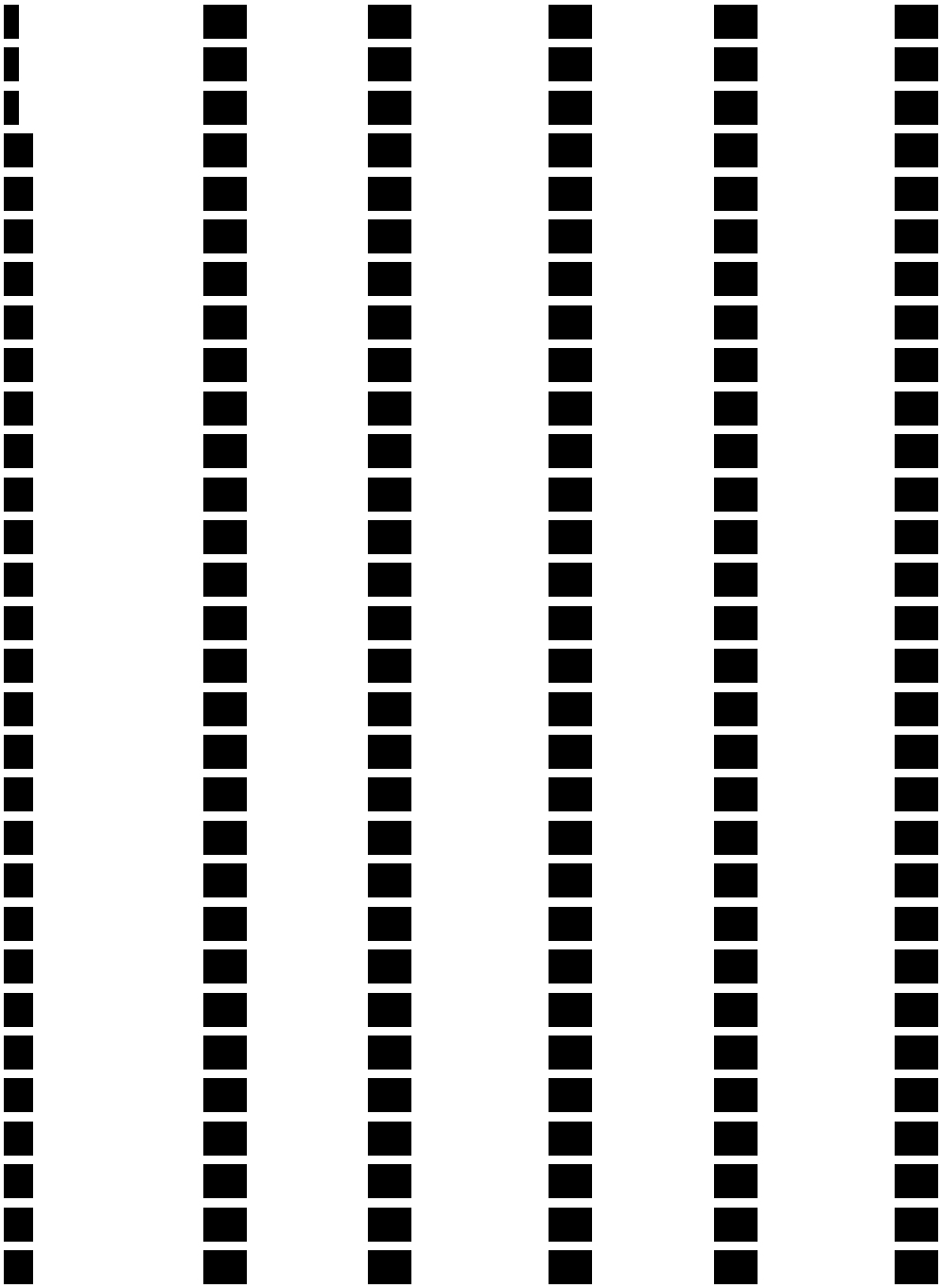
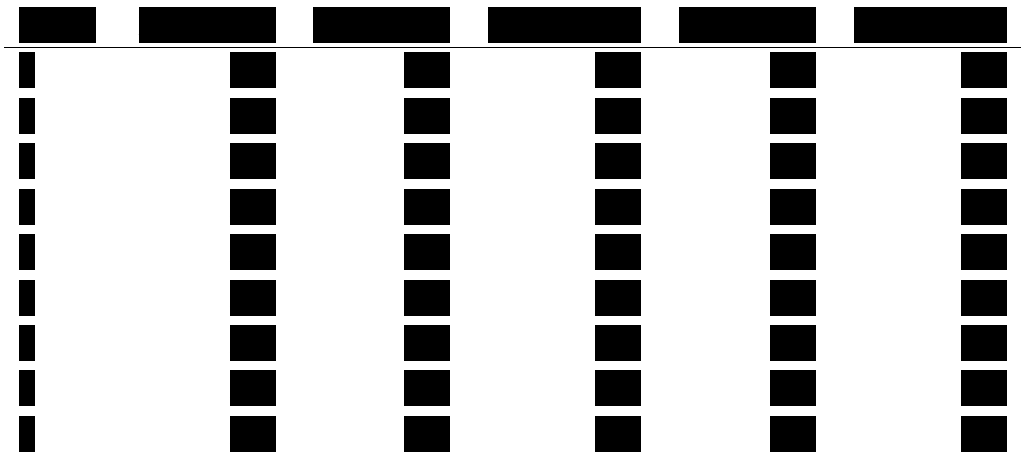


Table 108 PFS fractional polynomial model HR plot – random-effect model sorafenib



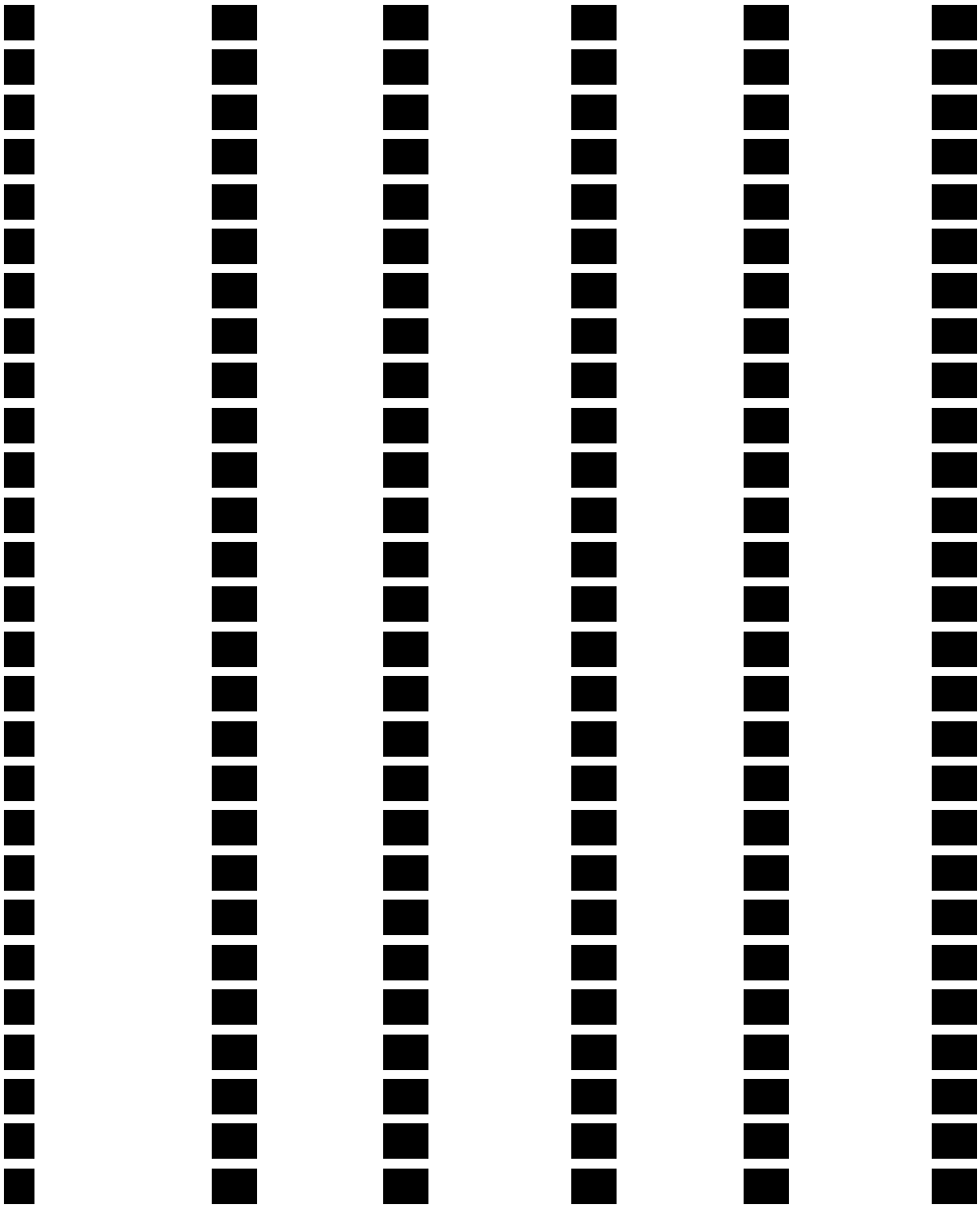
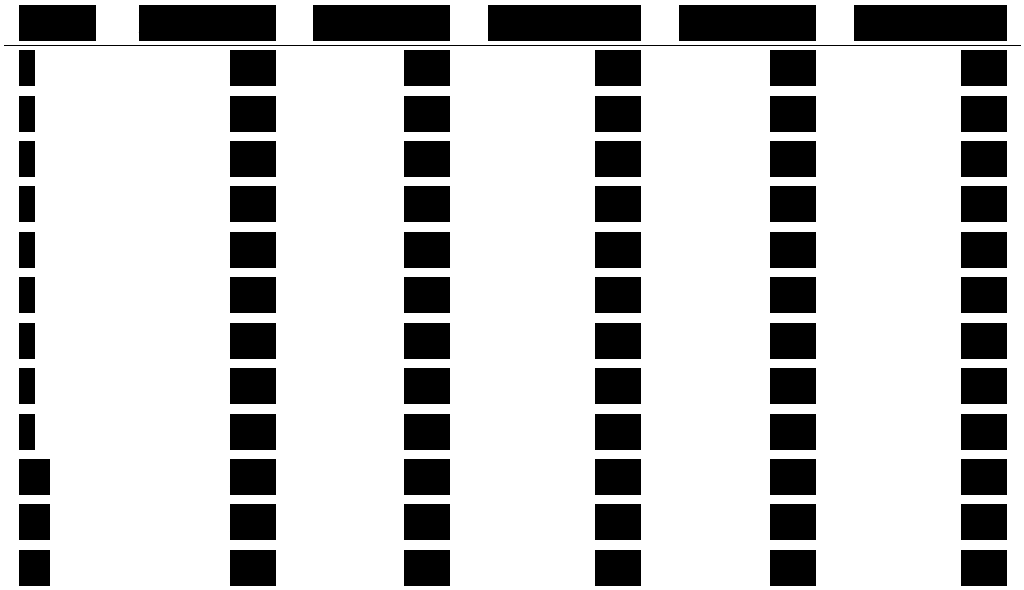


Table 109 PFS fractional polynomial model HR plot – random-effect model sunitinib



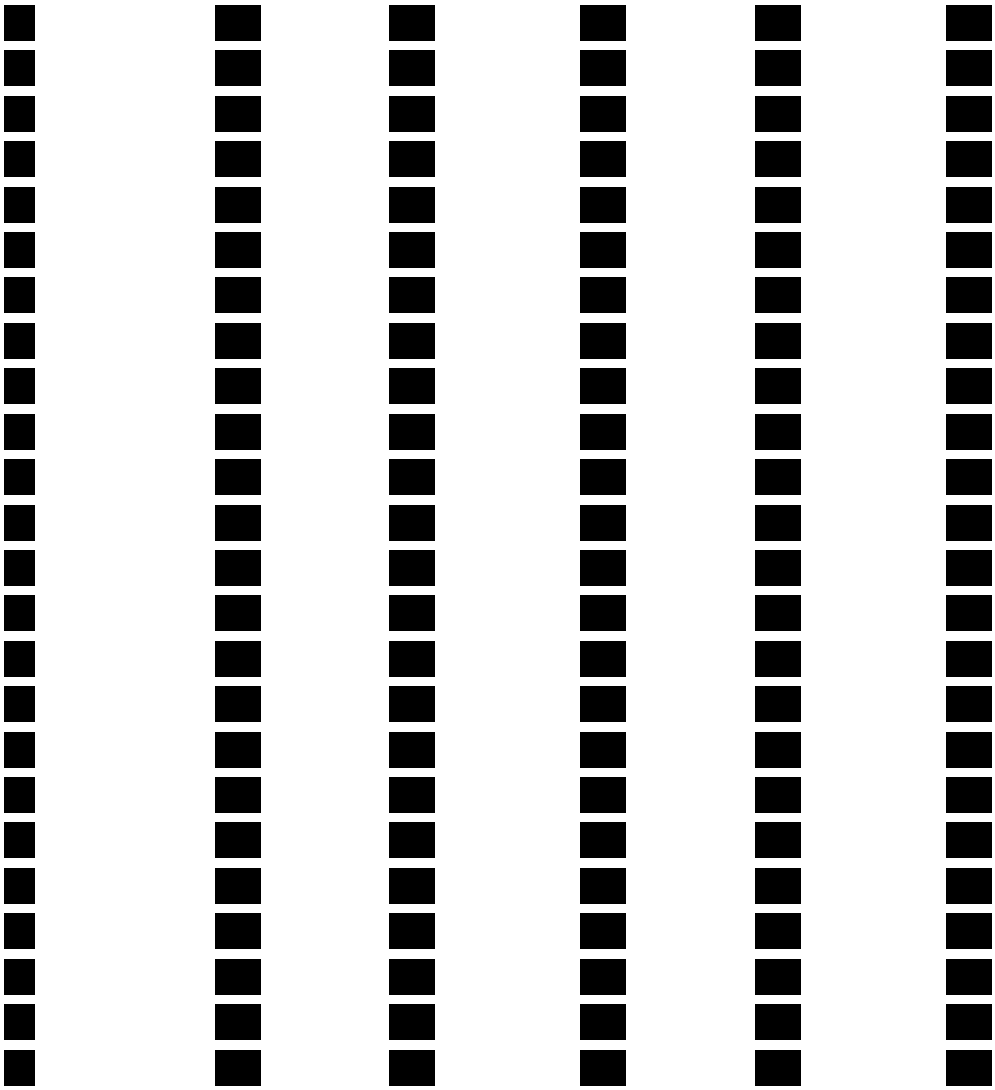


Table 110 PFS fractional polynomial model HR plot – random-effect model temsirolimus

[redacted]

Table 111 PFS fractional polynomial model HR plot – random-effect model tivozanib

[redacted]

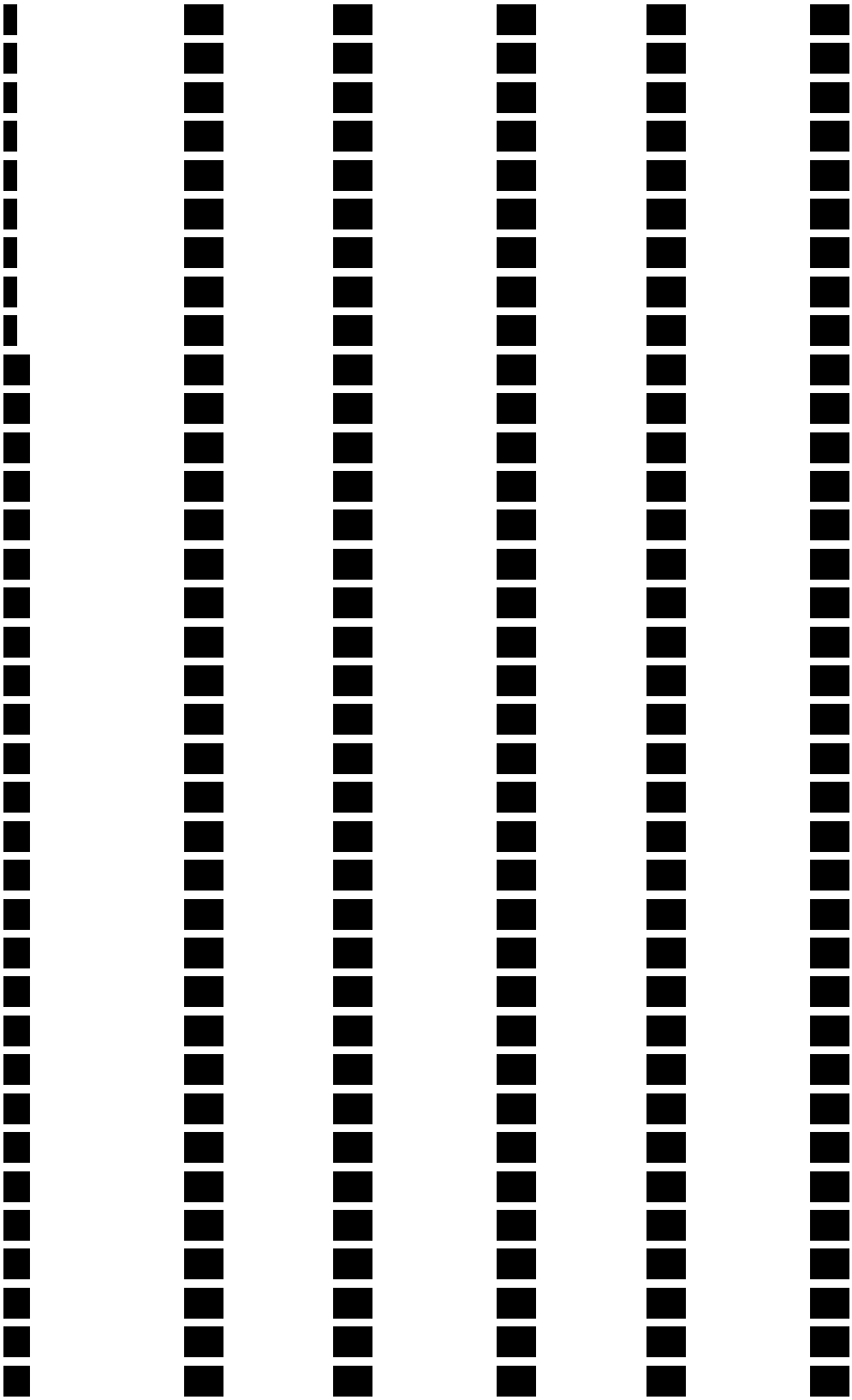


Table 112 Statistical fits – DIC fractional polynomial random-effect

FP	parameters	OS	PFS
1st	P0	5756.1	9561.5

1st	P1	5789.9	9582.4
1st	P05	5780.1	9584.0
1st	Pm1	5665.4	9347.8
1st	Pm05	5711.8	9476.3
2nd	Pm050	5614.8	8975.9
2nd	Pm10	5603.6	8899.8
2nd	Pm1m1	5598.0	8808.0
2nd	Pm105	5610.9	8965.7
2nd	Pm11	5620.0	9033.6

A27.

Priority question. For overall and progression-free survival, please provide graphs showing the CABOSUN and regenerated COMPARZ Kaplan–Meier graphs overlaid on the fitted survival curves (in appendix D) for each of the ITC parametric and fractional polynomial models.

Response: The re-generated data is provided in response to question A27. The plots with overlaid survival curves are shown from Figure 183 to Figure 214.

Figure 183. Fractional polynomial first order cabozantinib CABOSUN (OS)

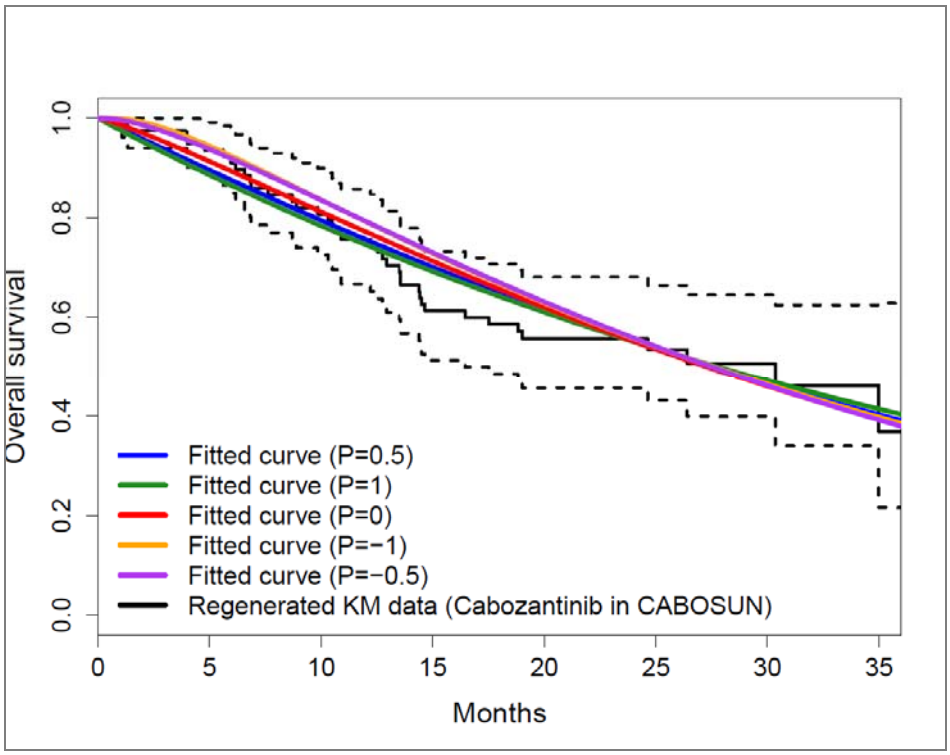


Figure 184. Fractional polynomial first order pazopanib COMPARZ (OS)

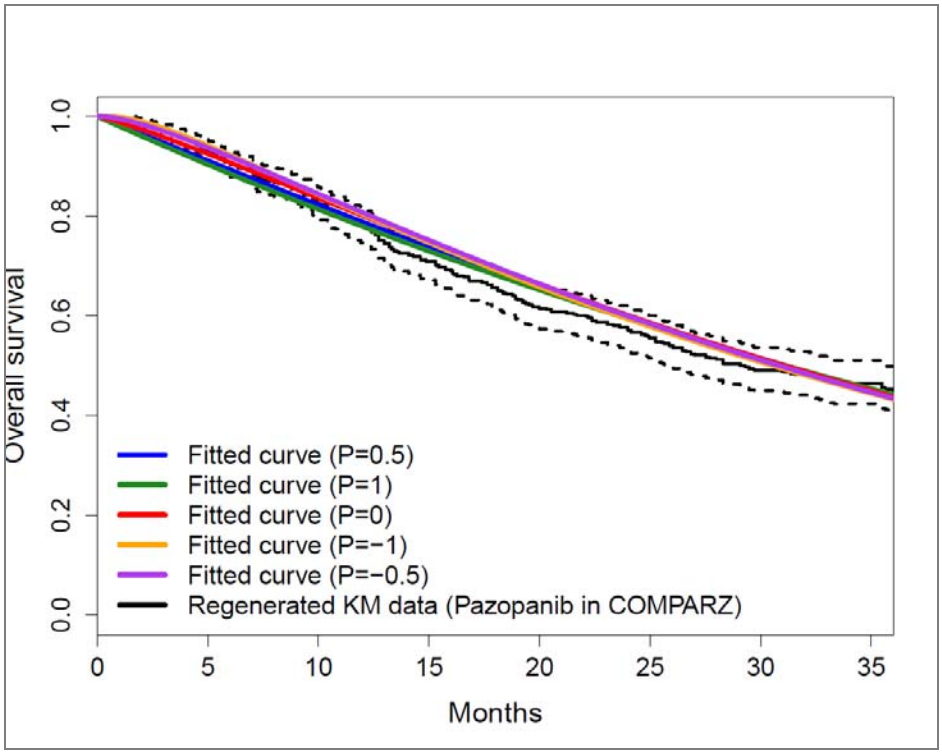


Figure 185. Fractional polynomial first order sunitinib CABOSUN (OS)

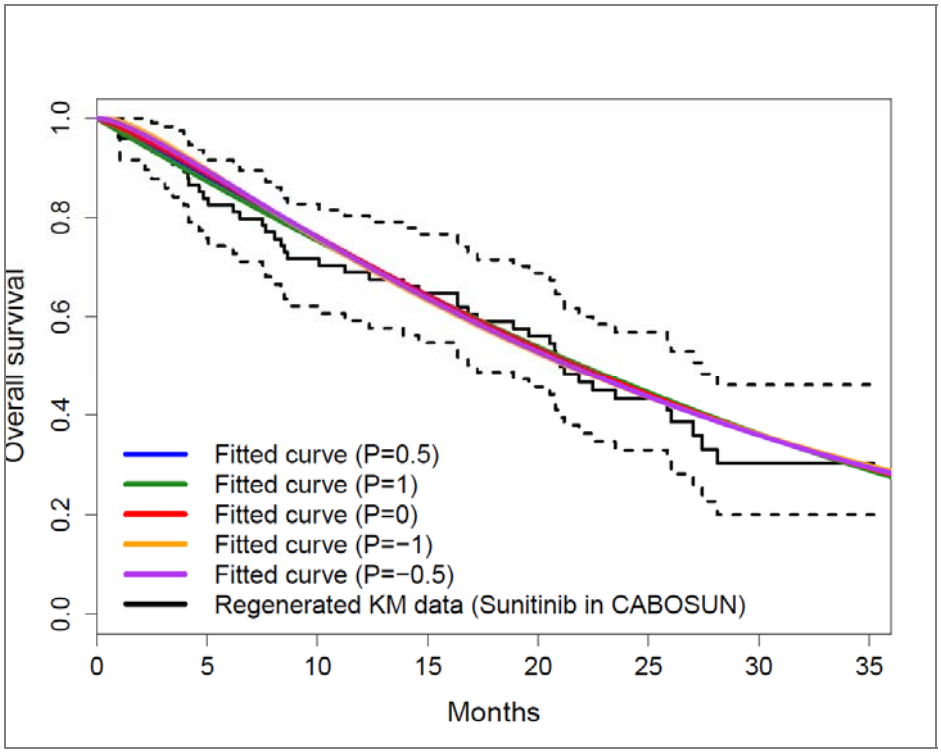


Figure 186. Fractional polynomial first order sunitinib COMPARZ (OS)

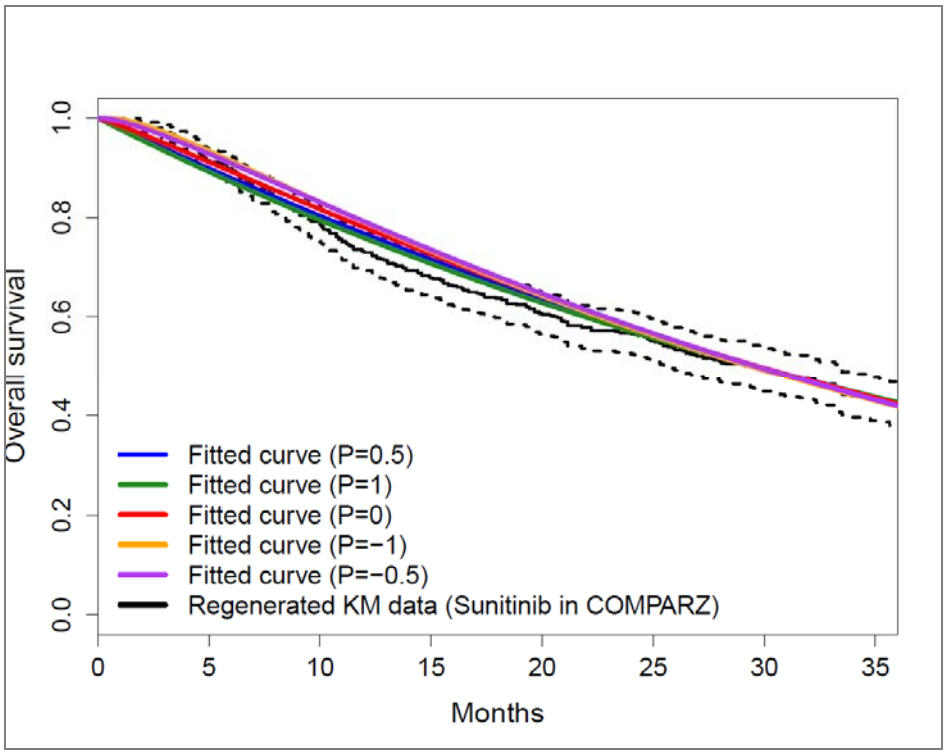


Figure 187. Fractional polynomial second order cabozantinib CABOSUN (OS)

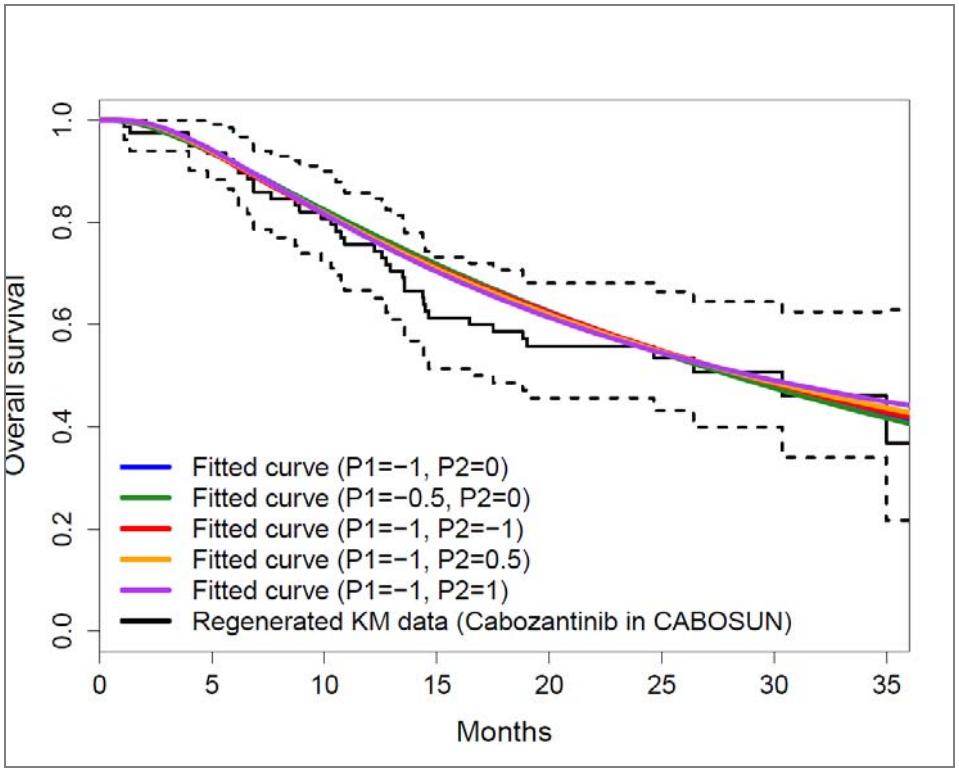


Figure 188. Fractional polynomial second order pazopanib COMPARZ (OS)

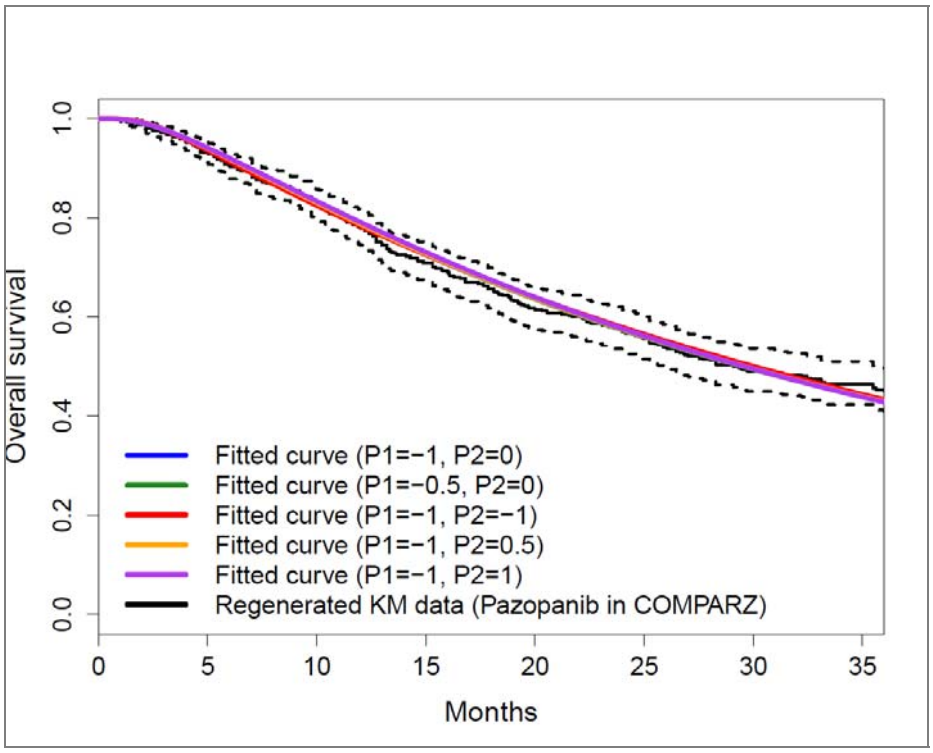


Figure 189. Fractional polynomial second order sunitinib CABOSUN (OS)

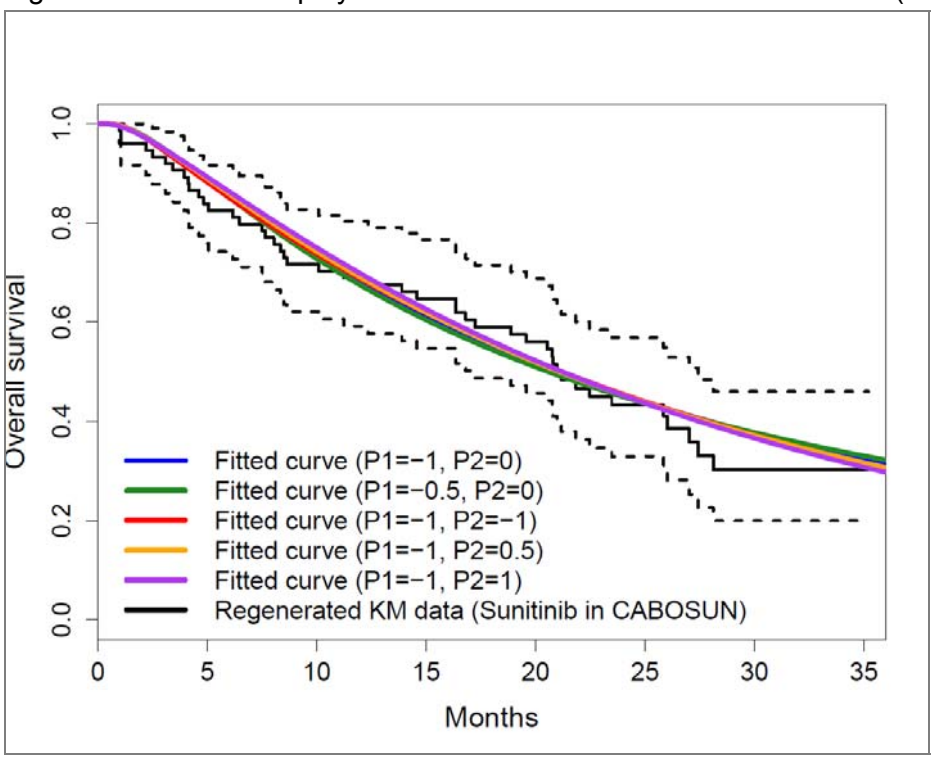


Figure 190. Fractional polynomial second order sunitinib COMPARZ (OS)

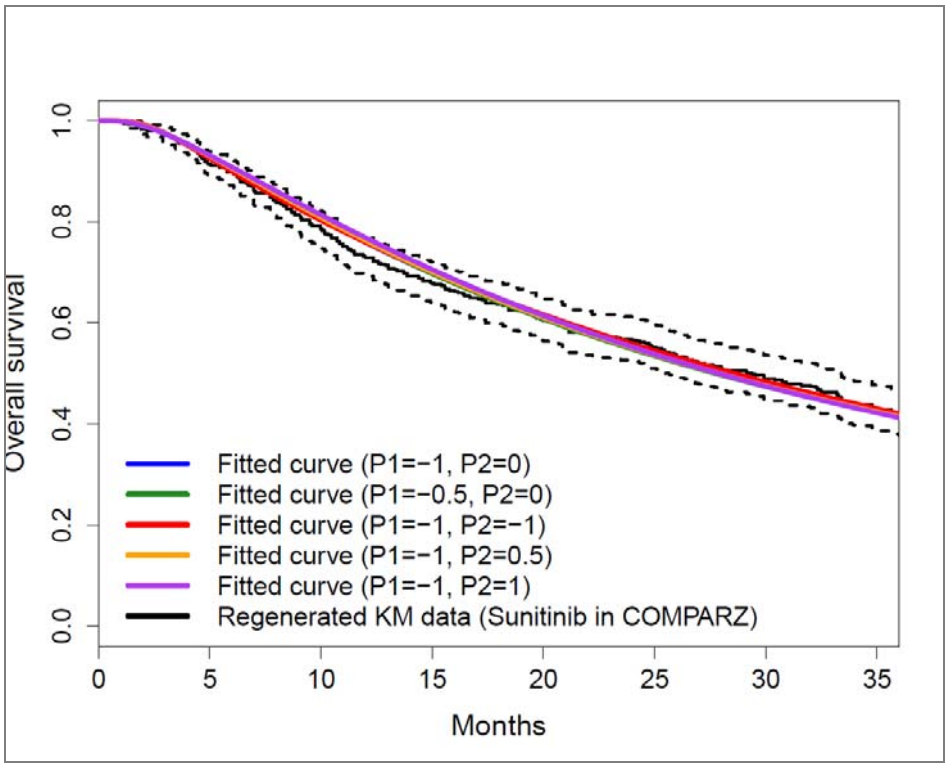


Figure 191. Fractional polynomial first order cabozantinib CABOSUN (PFS)

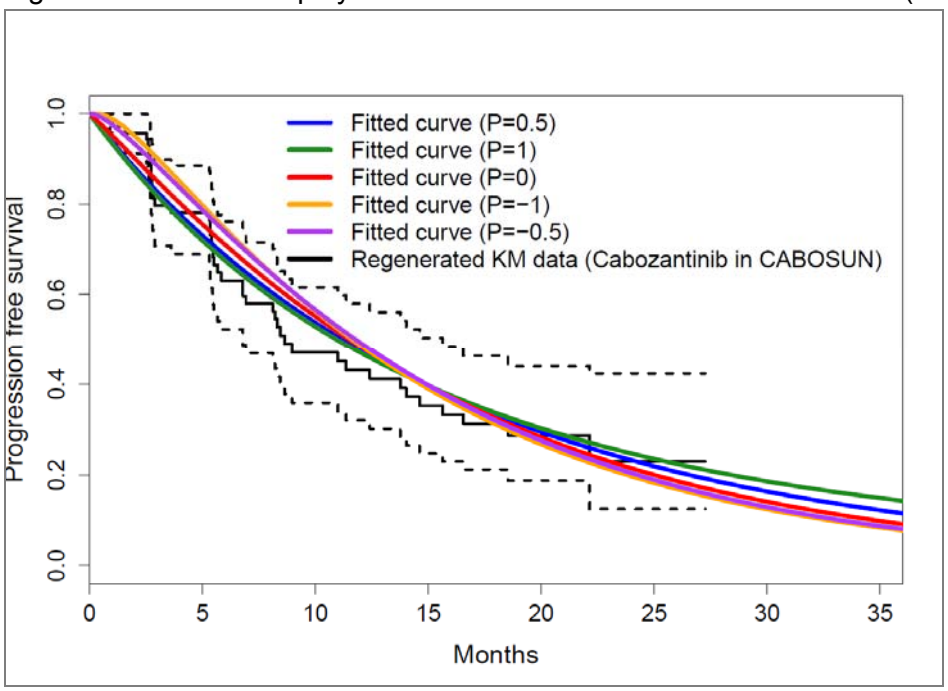


Figure 192. Fractional polynomial first order pazopanib COMPARZ (PFS)

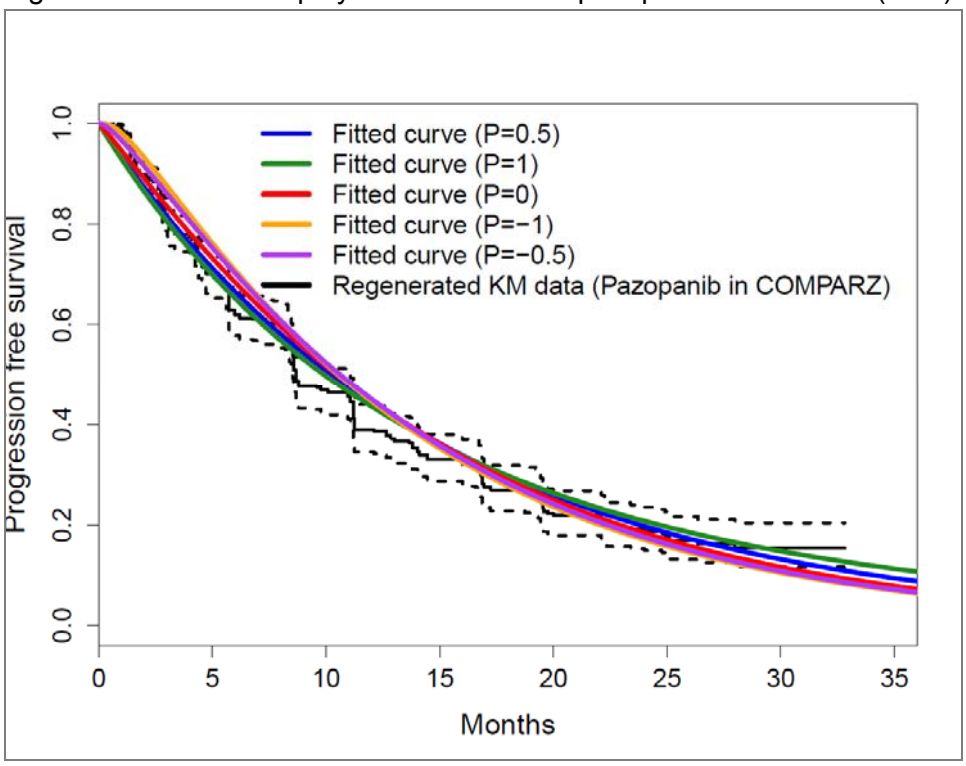


Figure 193. Fractional polynomial first order sunitinib CABOSUN (PFS)

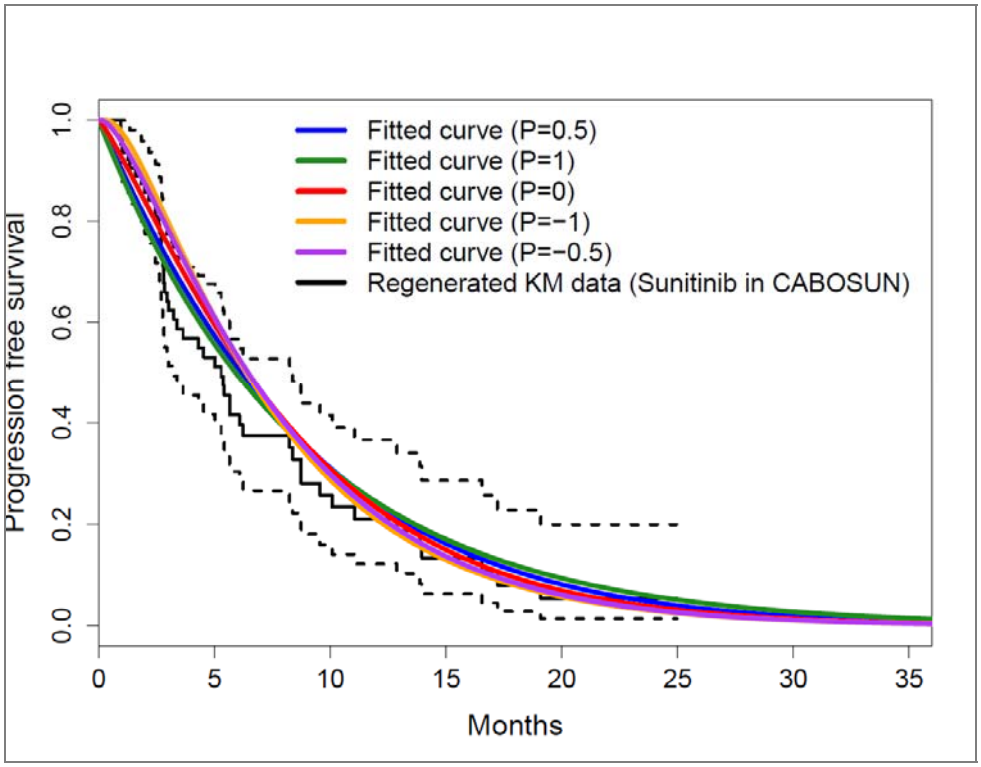


Figure 194. Fractional polynomial first order sunitinib COMPARZ (PFS)

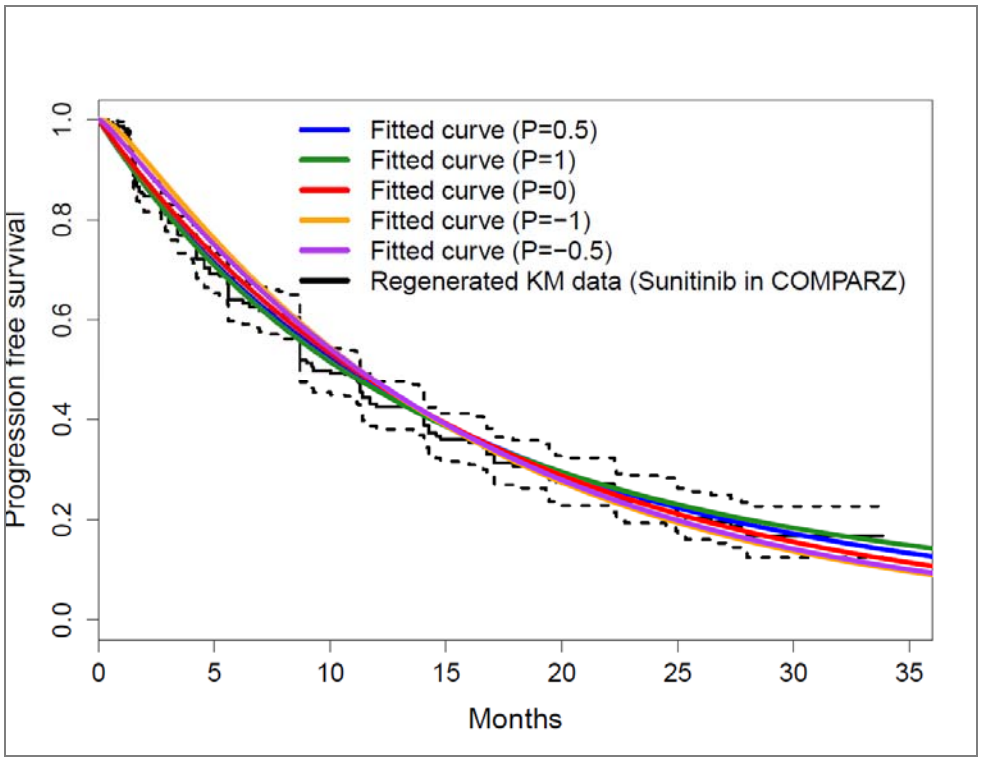


Figure 195. Fractional polynomial second order cabozantinib CABOSUN (PFS)

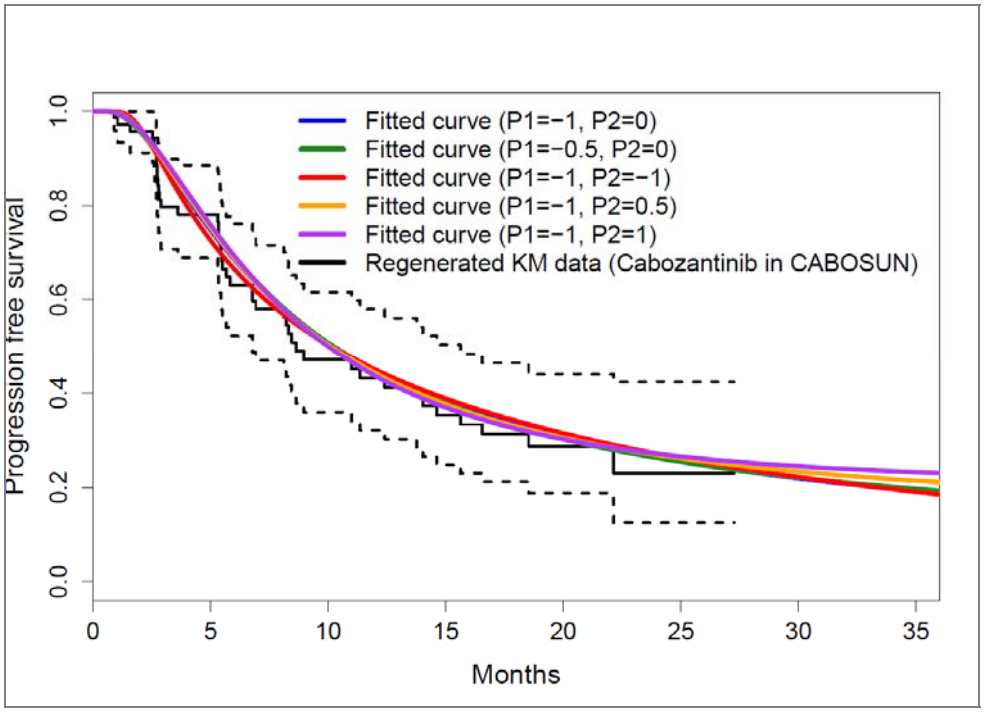


Figure 196. Fractional polynomial second order pazopanib COMPARZ (PFS)

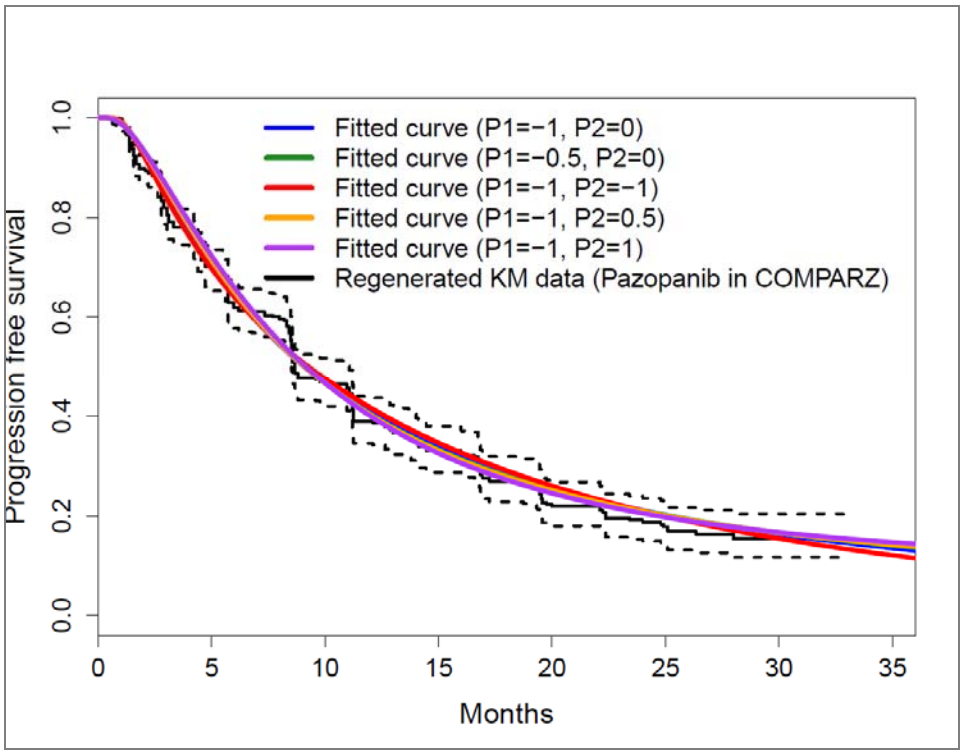


Figure 197. Fractional polynomial second order sunitinib CABOSUN (PFS)

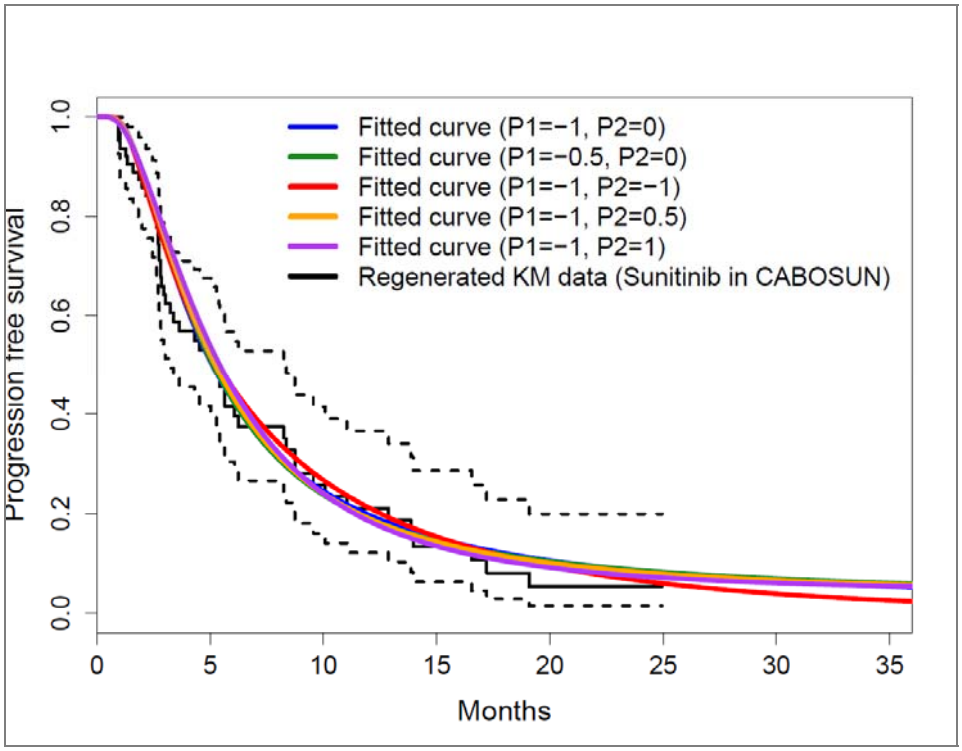


Figure 198. Fractional polynomial second order sunitinib COMPARZ (PFS)

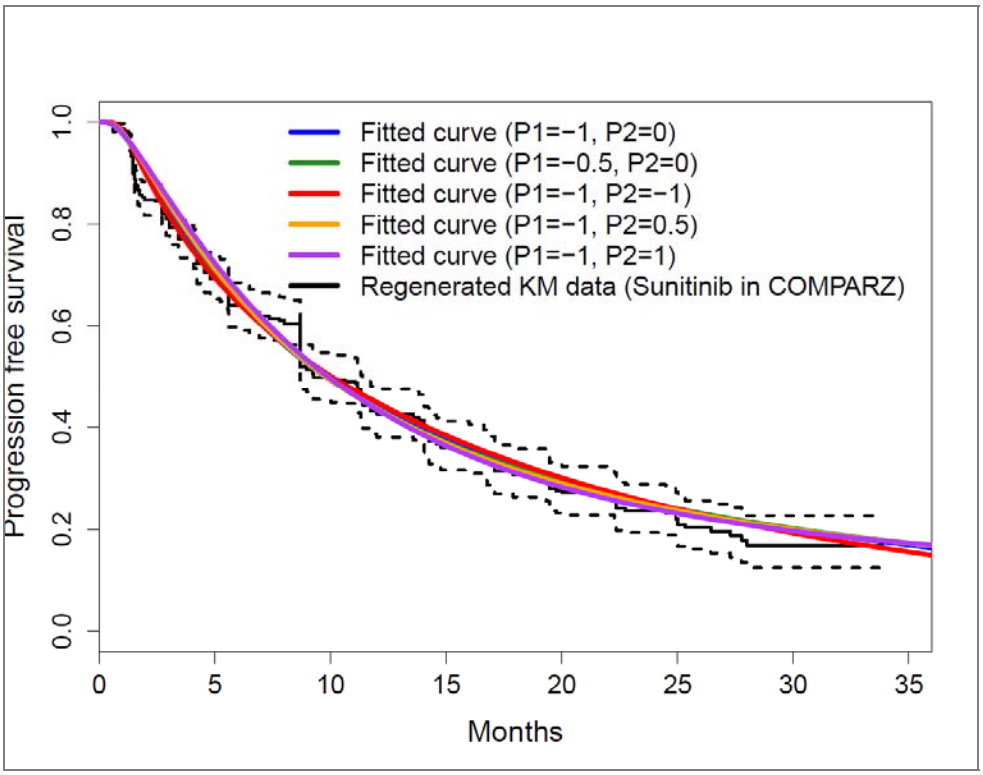


Figure 199. Ouwens et al. Fixed effect cabozantinib CABOSUN (OS)

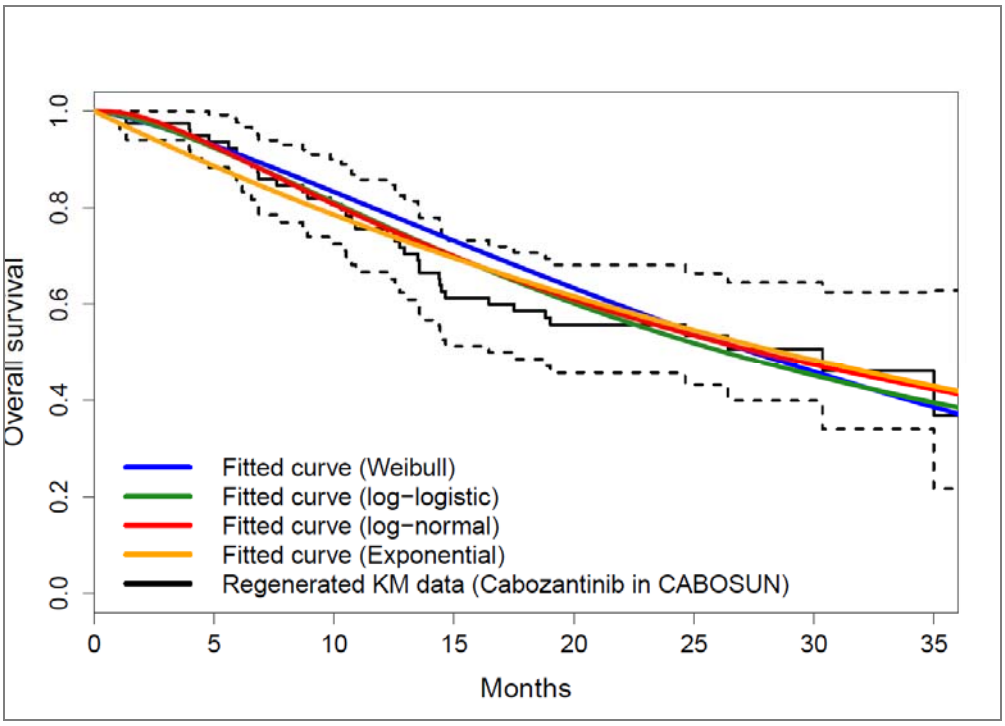


Figure 200. Ouwens et al. Fixed effect pazopanib COMPARZ (OS)

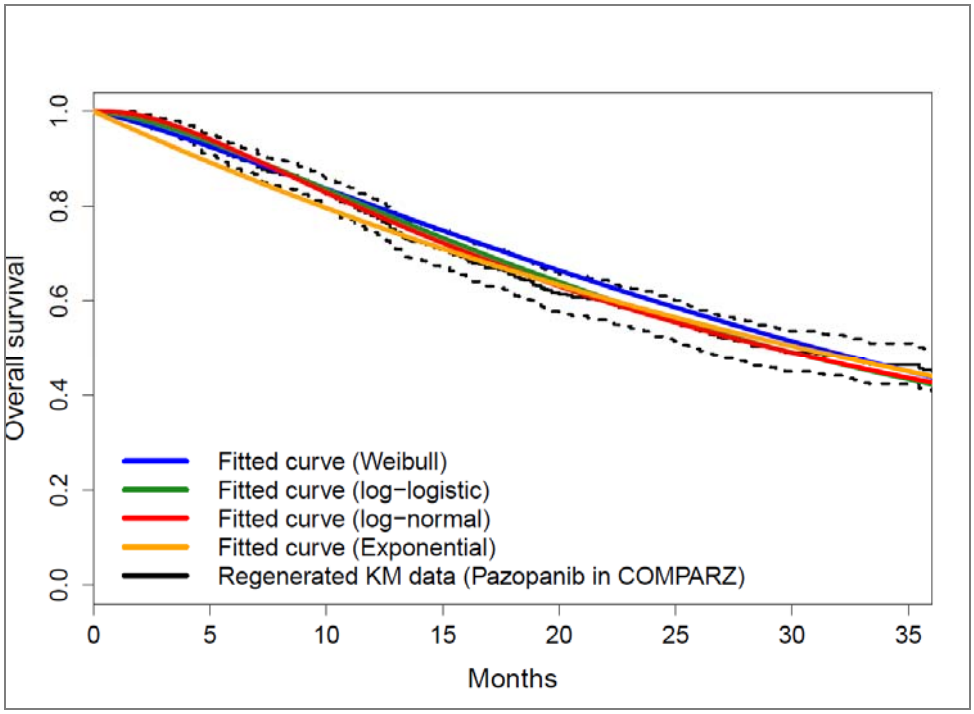


Figure 201. Ouwens et al. Fixed effect sunitinib CABOSUN (OS)

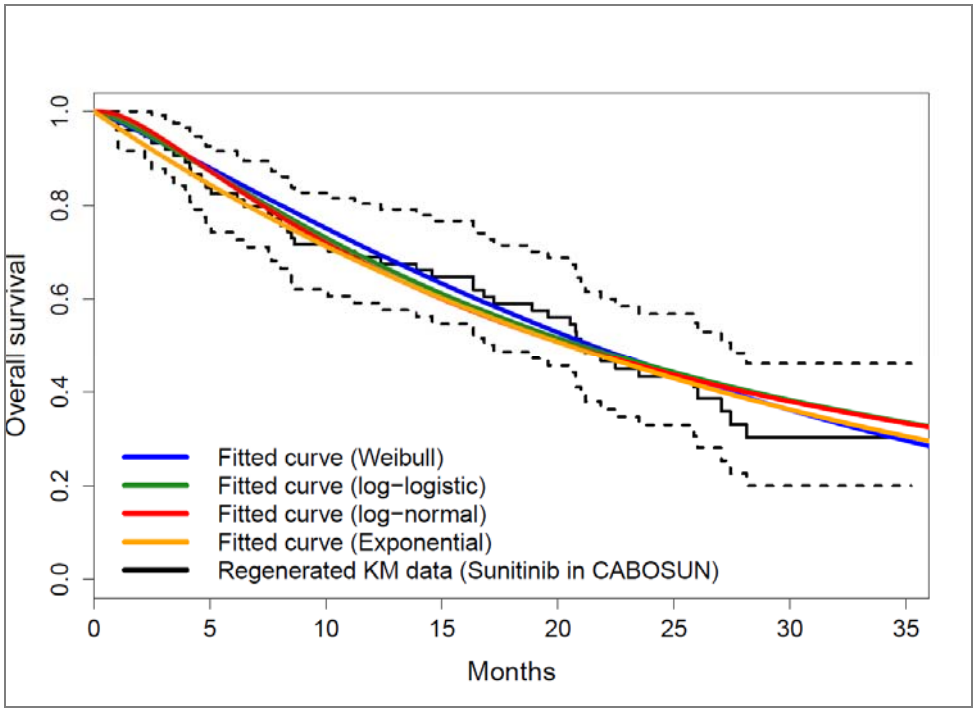


Figure 202. Ouwens et al. Fixed effect sunitinib COMPARZ (OS)

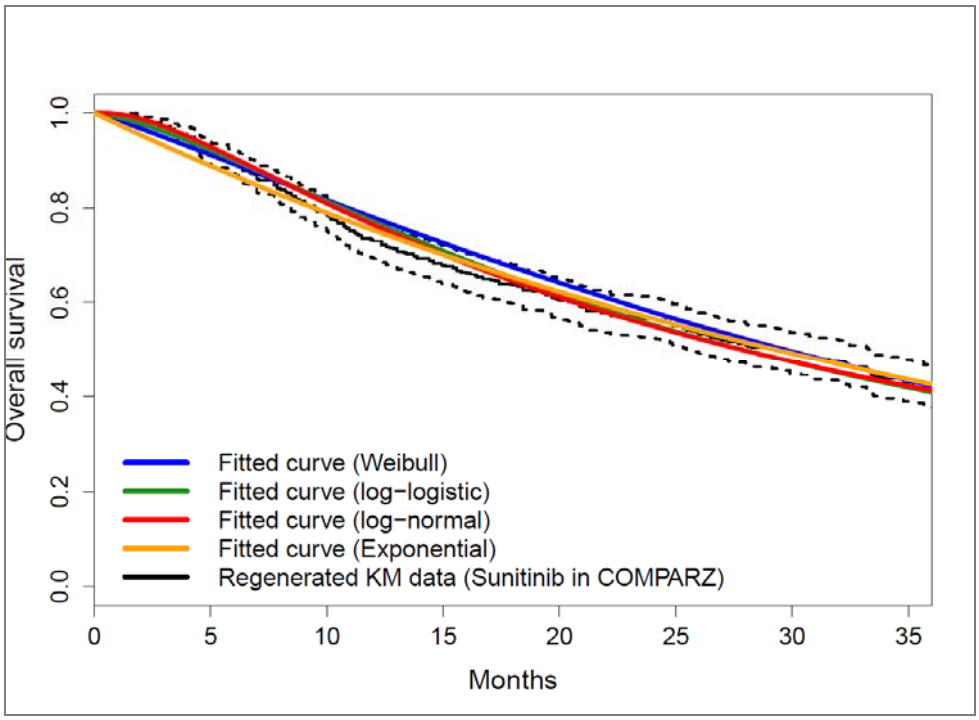


Figure 203. Ouwens et al. Random effect cabozantinib CABOSUN (OS)

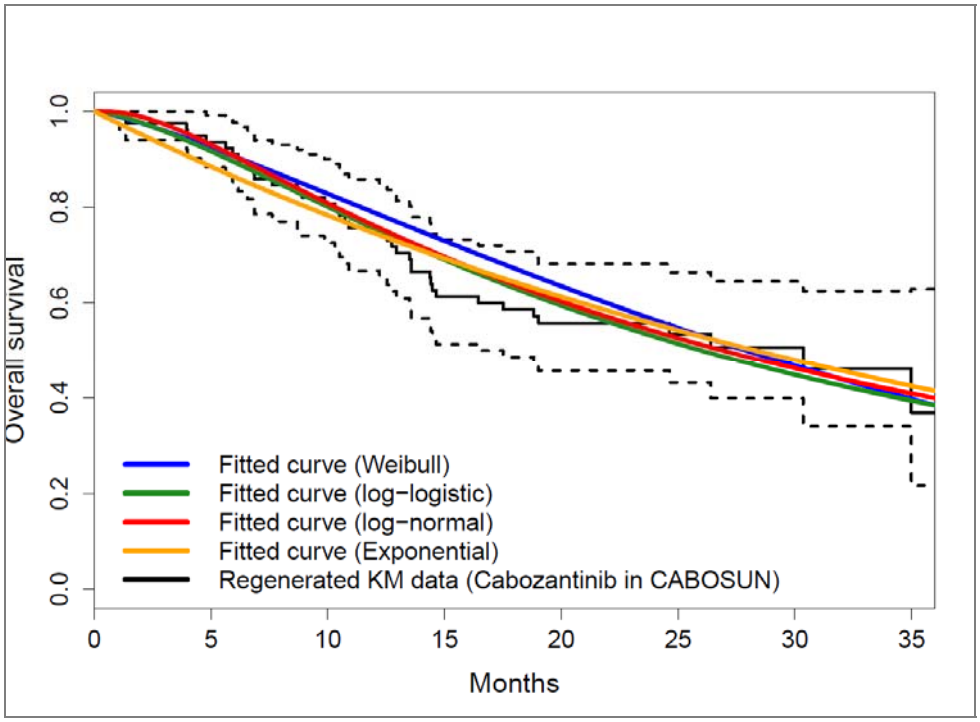


Figure 204. Ouwens et al. Random effect pazopanib COMPARZ (OS)

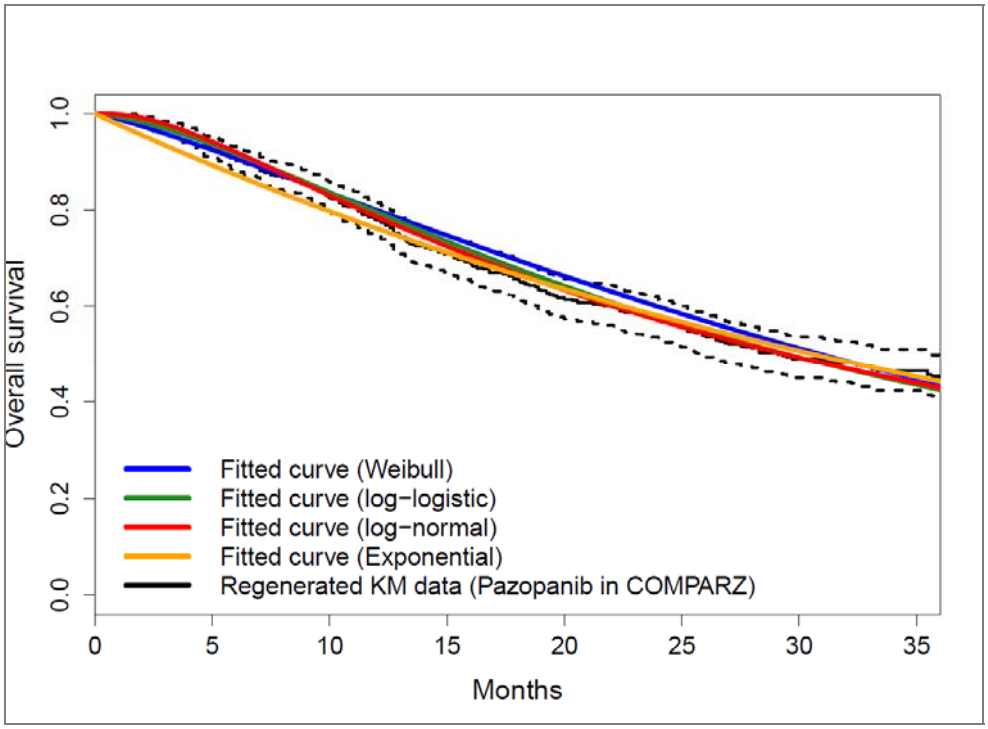


Figure 205. Ouwens et al. Random effect sunitinib CABOSUN (OS)

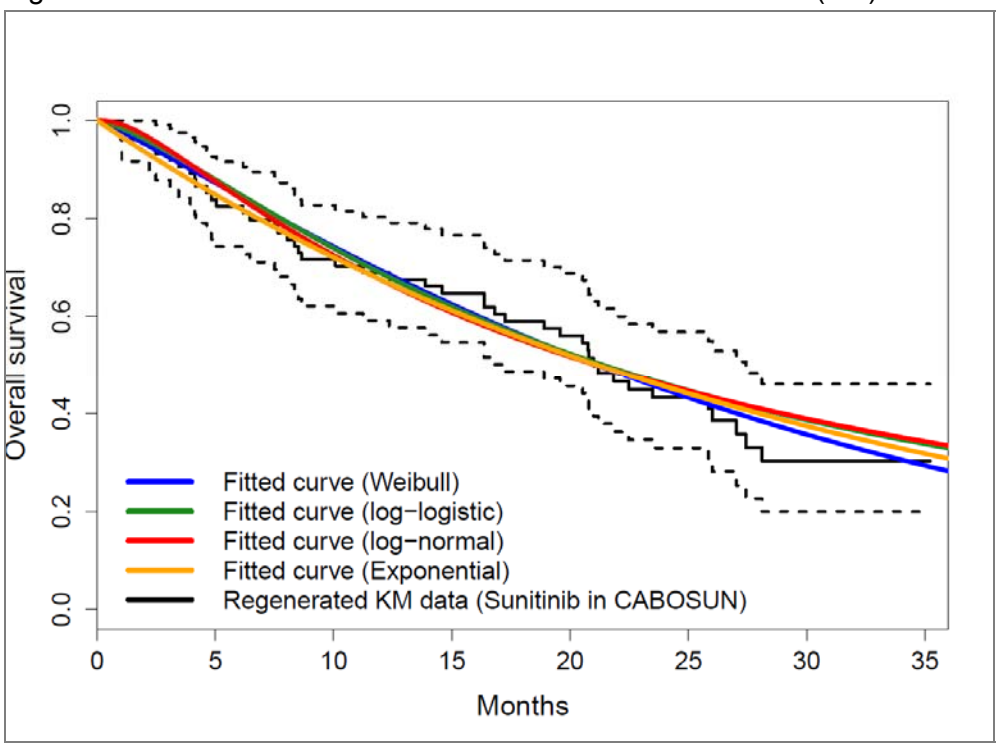


Figure 206. Ouwens et al. Random effect sunitinib COMPARZ (OS)

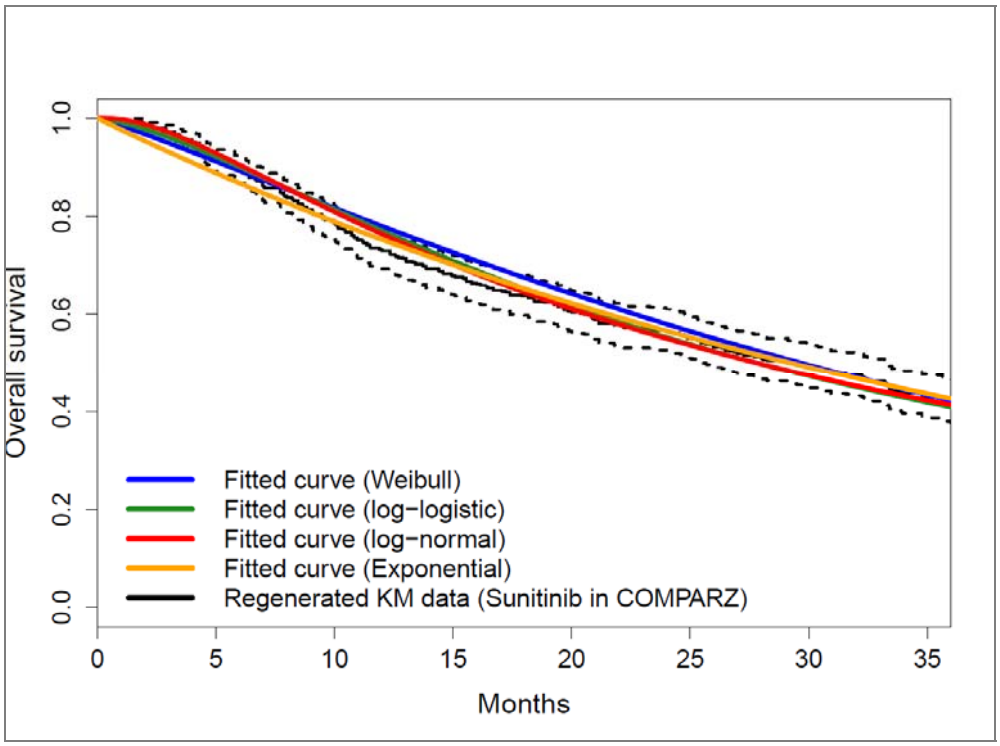


Figure 207. Ouwens et al. Fixed effect cabozantinib CABOSUN (PFS)

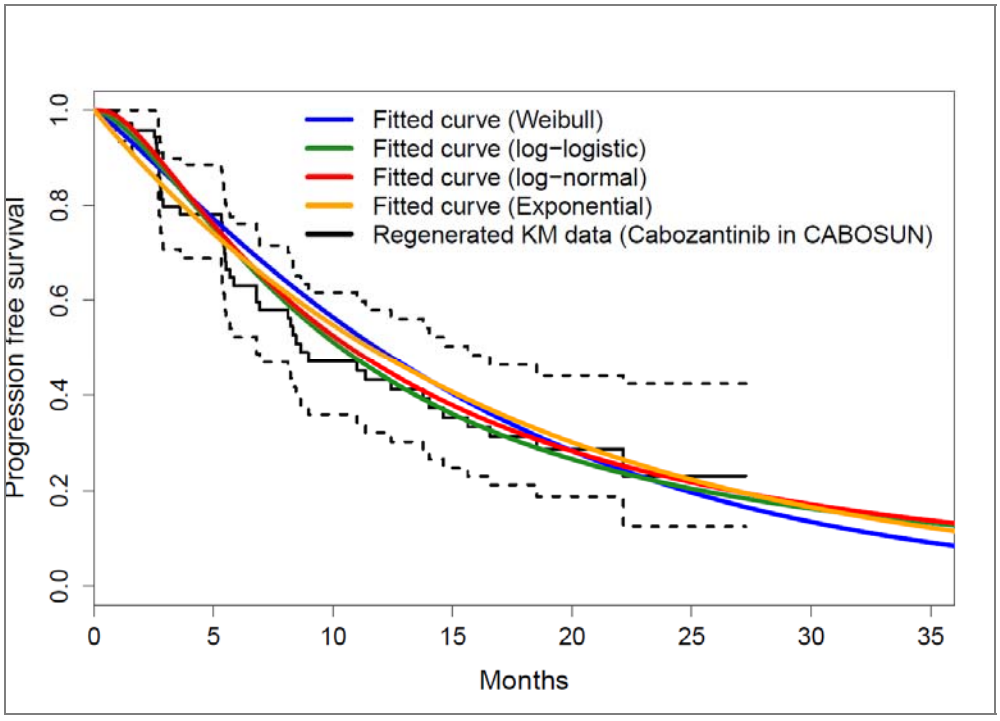


Figure 208. Ouwens et al. Fixed effect pazopanib COMPARZ (PFS)

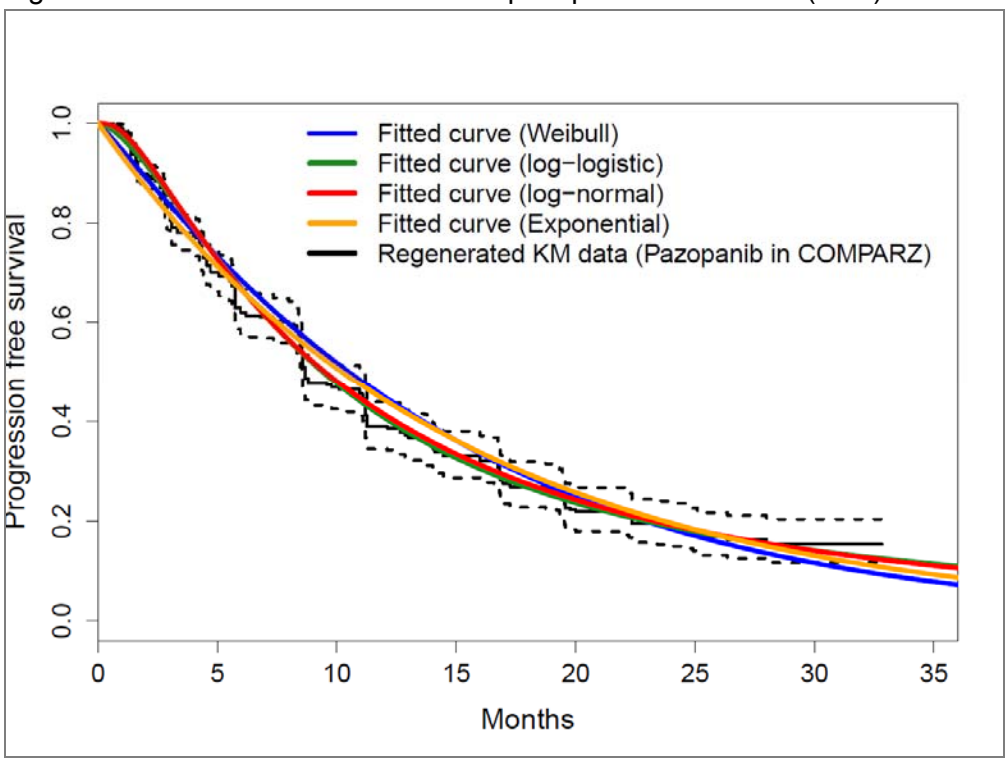


Figure 209. Ouwens et al. Fixed effect sunitinib CABOSUN (PFS)

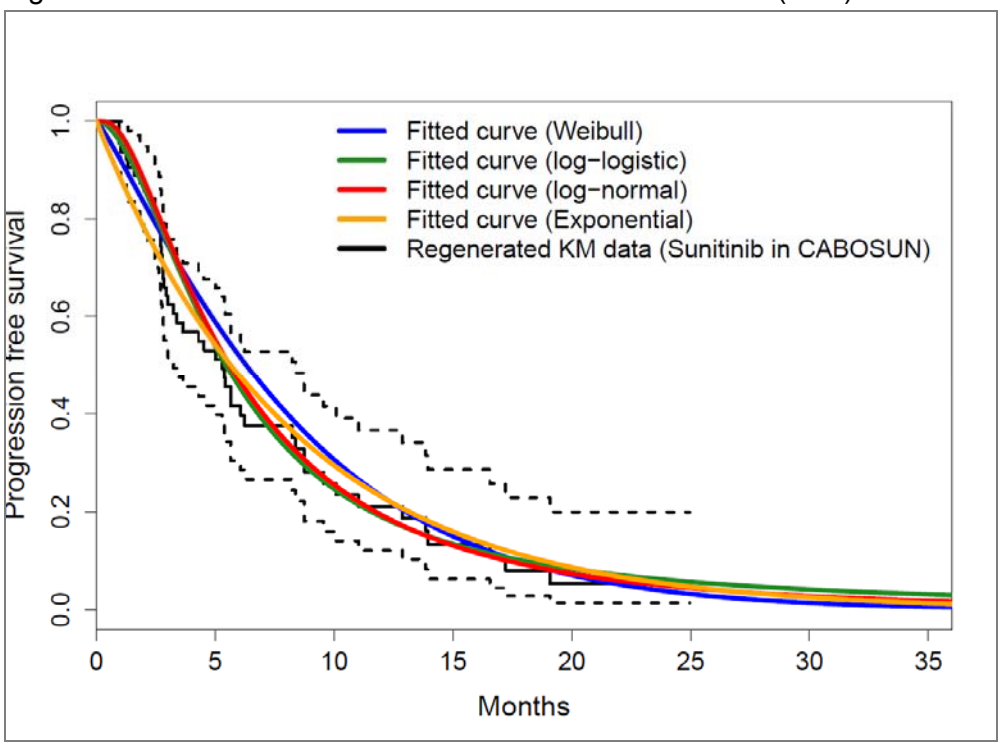


Figure 210. Ouwens et al. Fixed effect sunitinib COMPARZ (PFS)

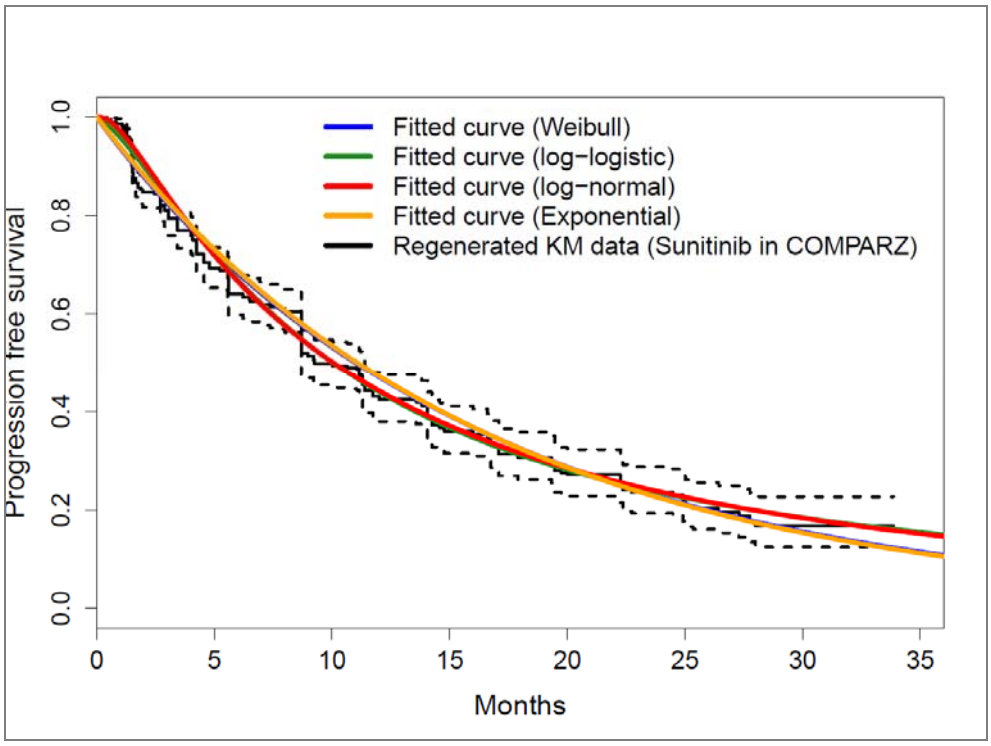


Figure 211. Ouwens et al. Random effect cabozantinib CABOSUN (PFS)

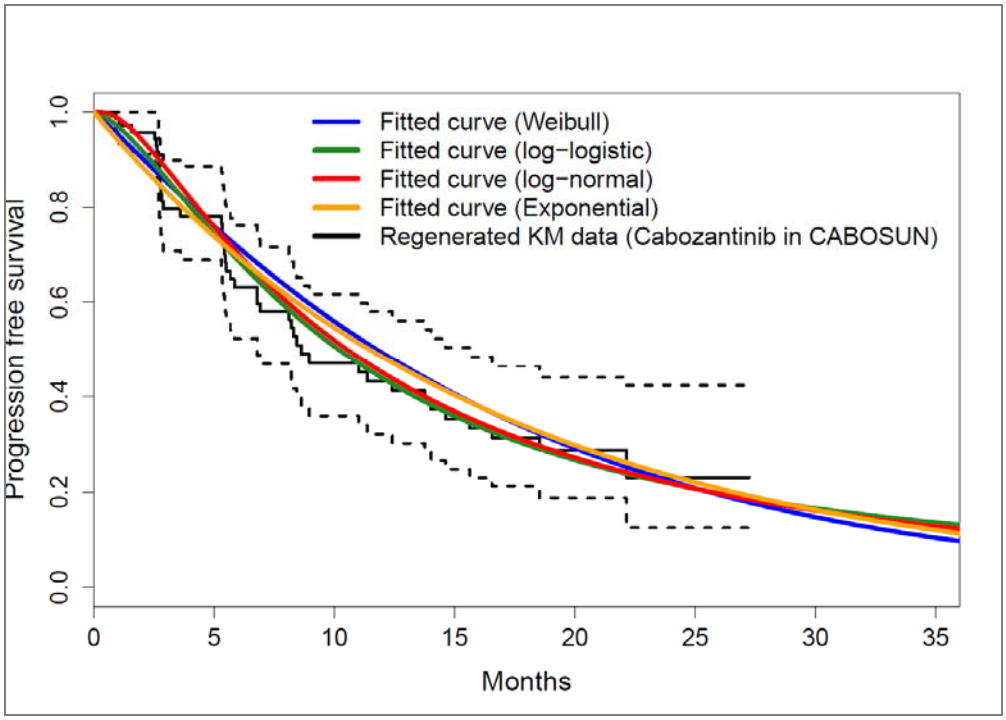


Figure 212. Ouwens et al. Random effect pazopanib COMPARZ (PFS)

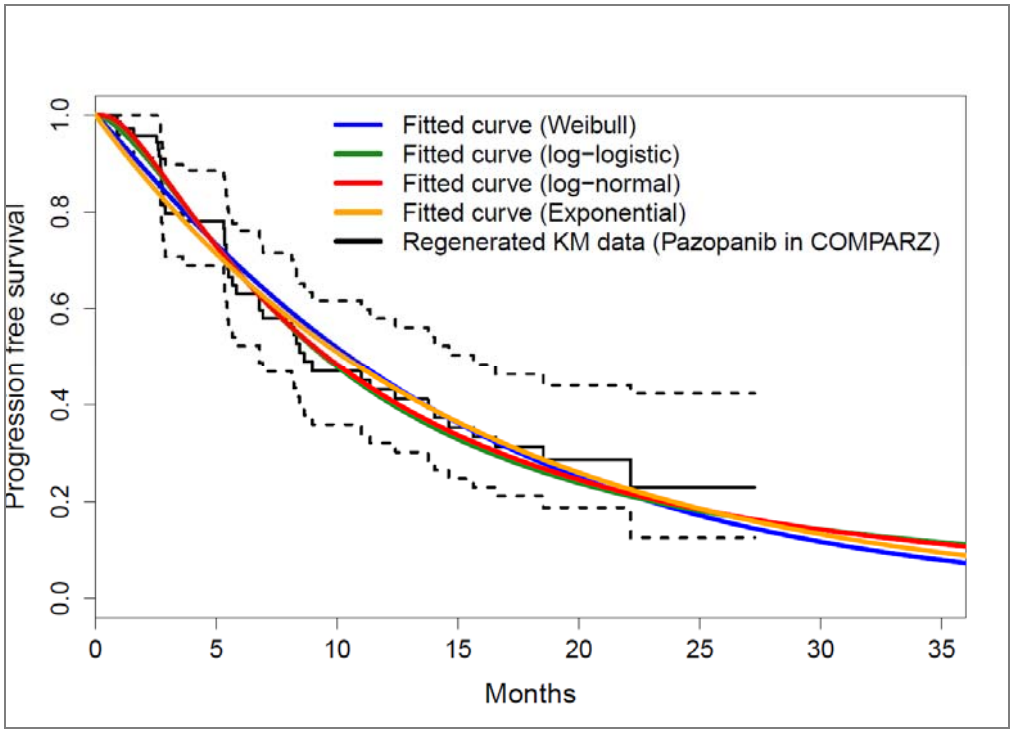


Figure 213. Ouwens et al. Random effect sunitinib CABOSUN (PFS)

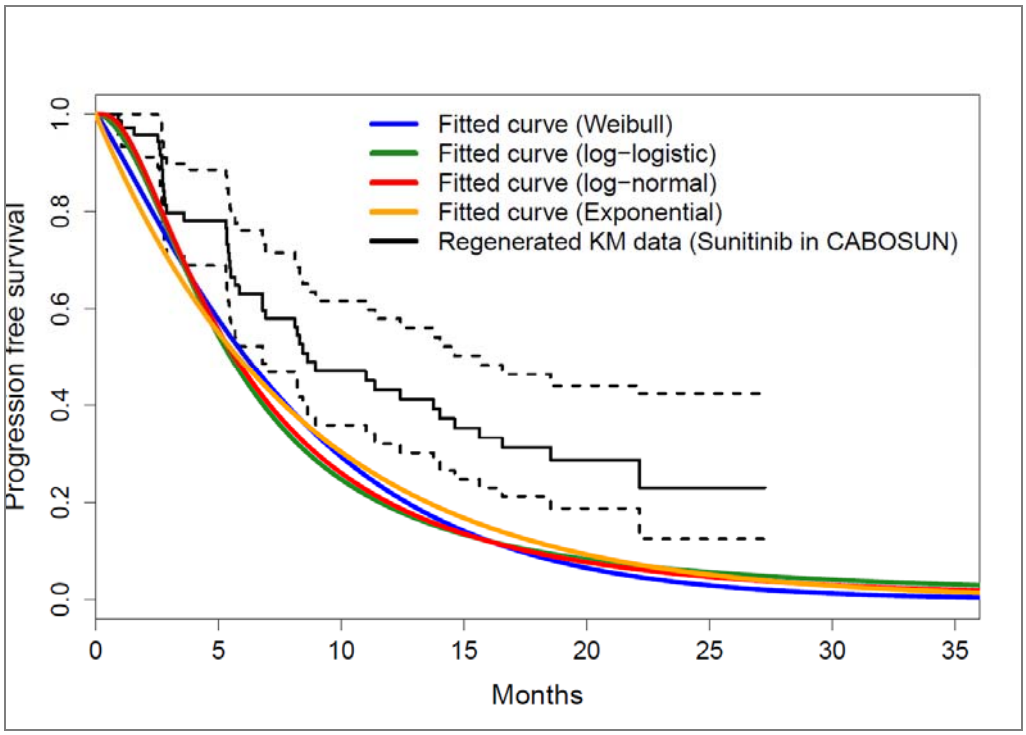
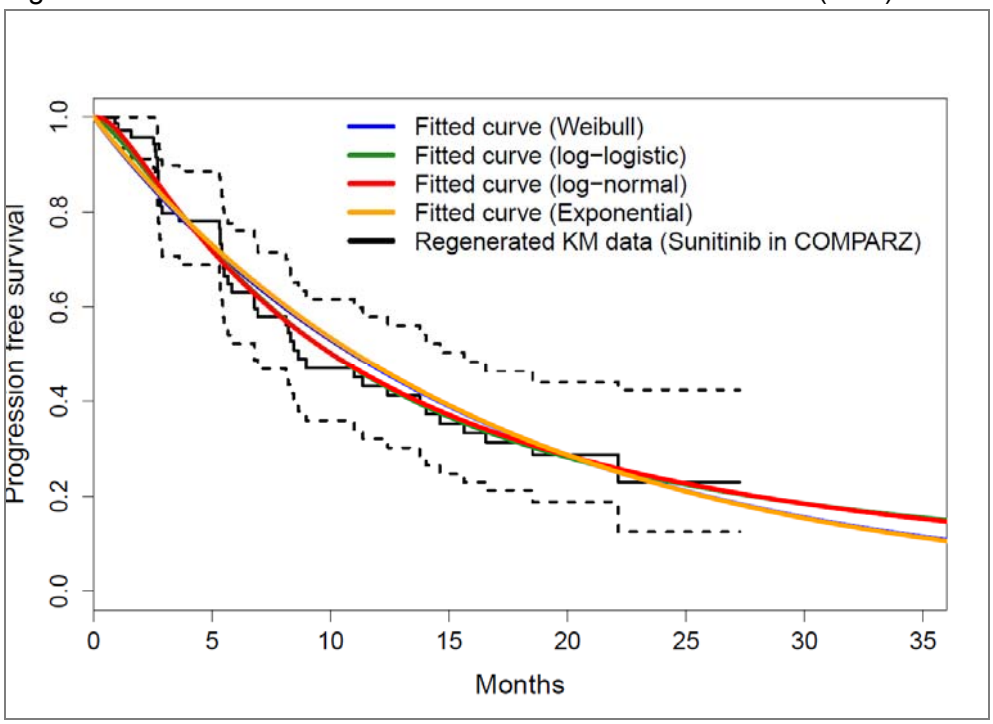


Figure 214. Ouwens et al. Random effect sunitinib COMPARZ (PFS)



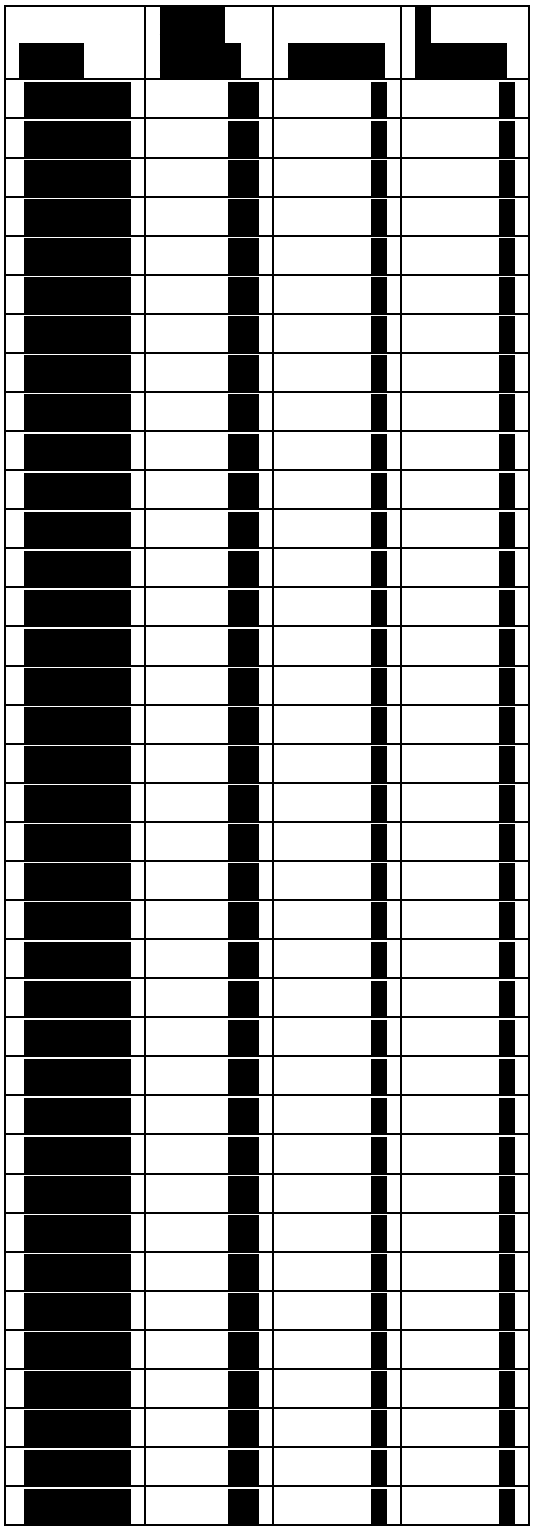
Section B: Clarification on cost-effectiveness data

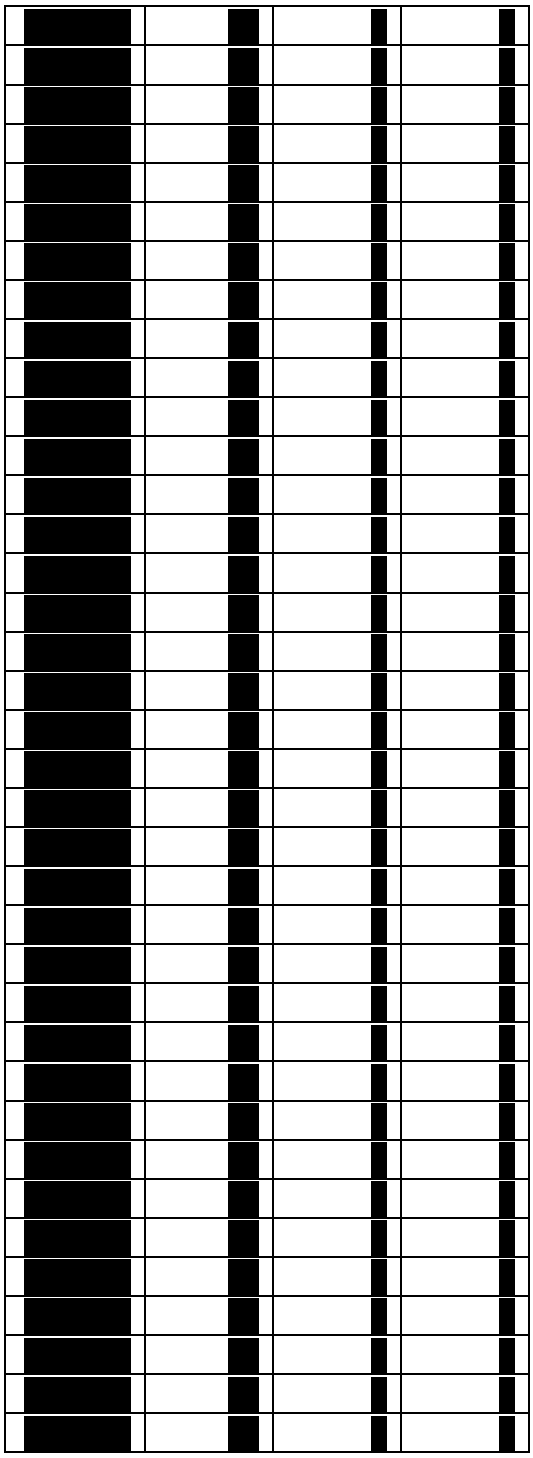
B1.

Priority question. Please provide Kaplan–Meier data from CABOSUN and the regenerated Kaplan–Meier data from COMPARZ that you used to fit the parametric and fractional polynomial overall and progression-free survival models.

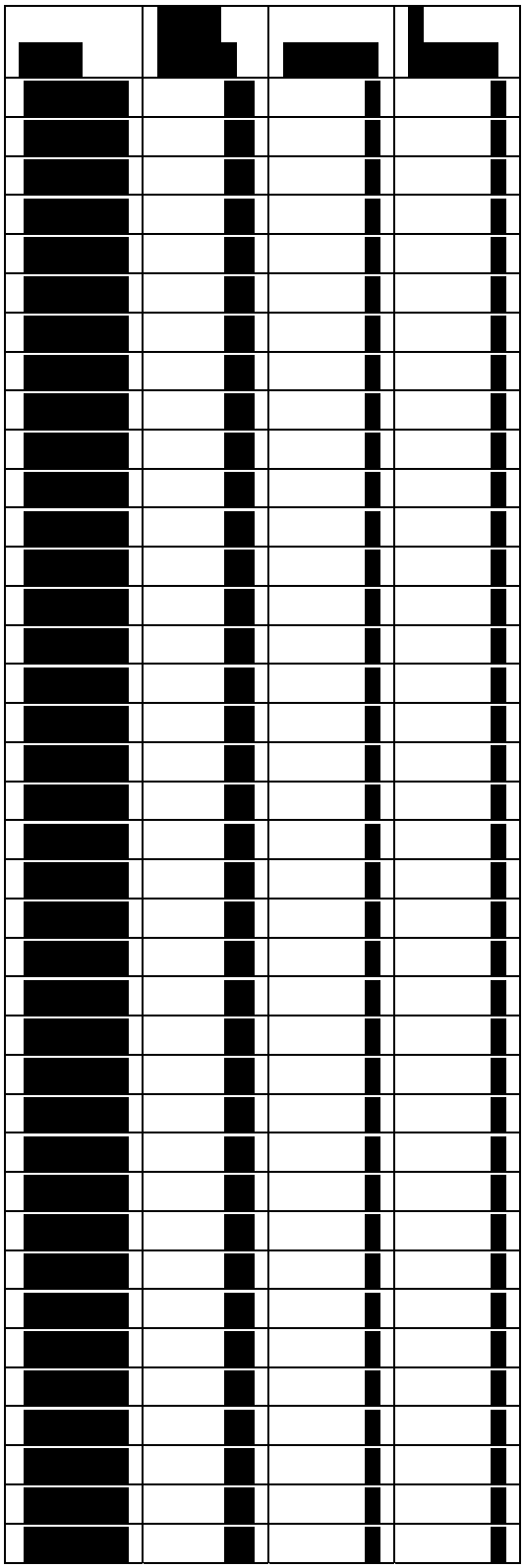
Response: Please find the requested data below.

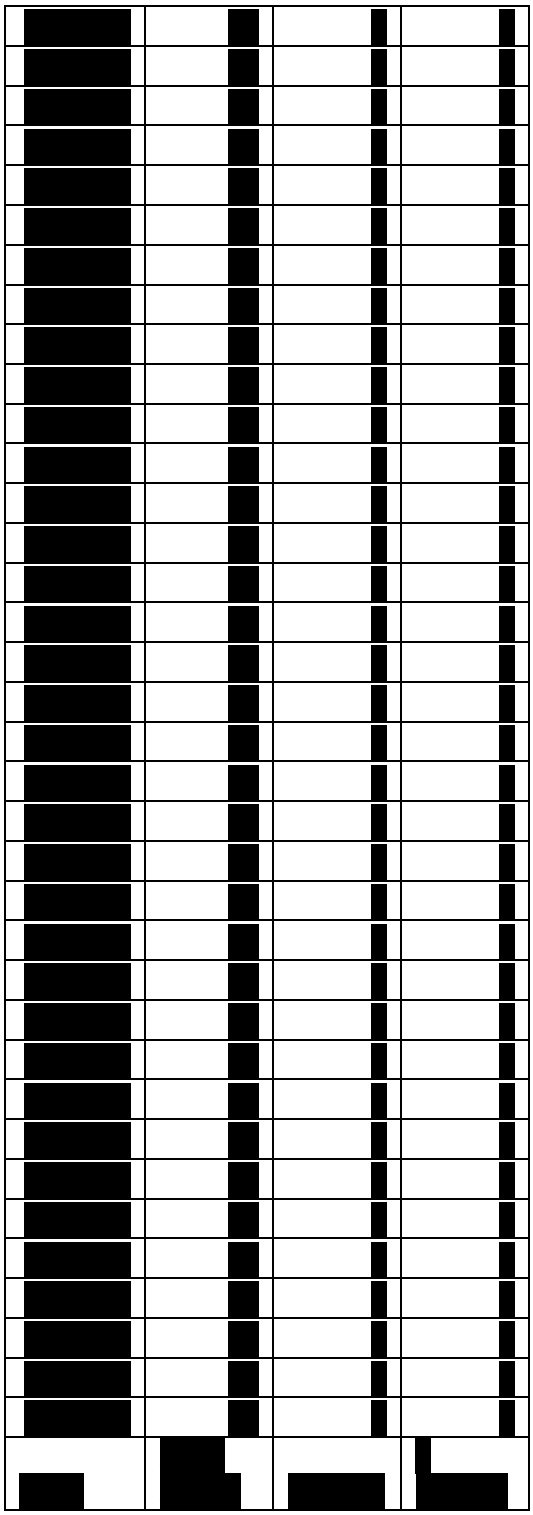
CABOSUN cabozantinib OS KM data



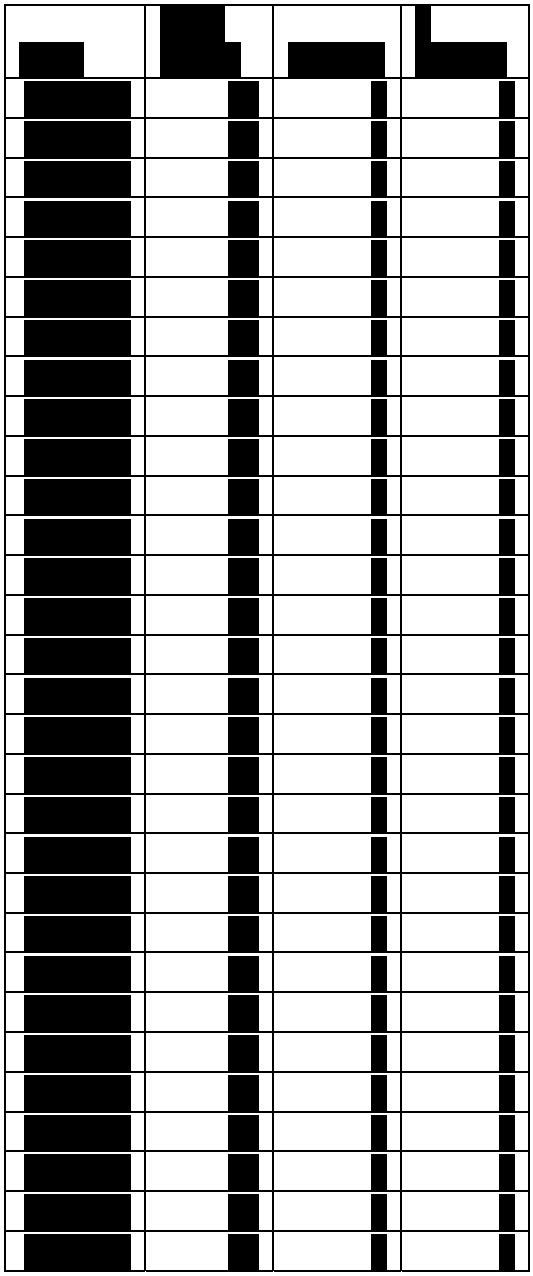


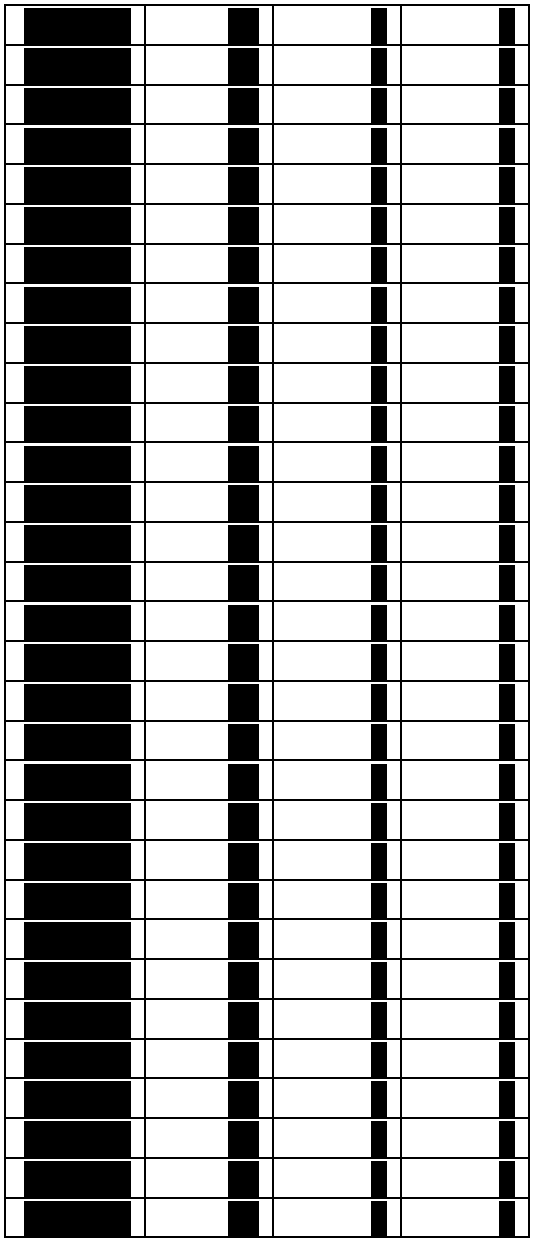
CABOSUN sunitinib OS KM data:



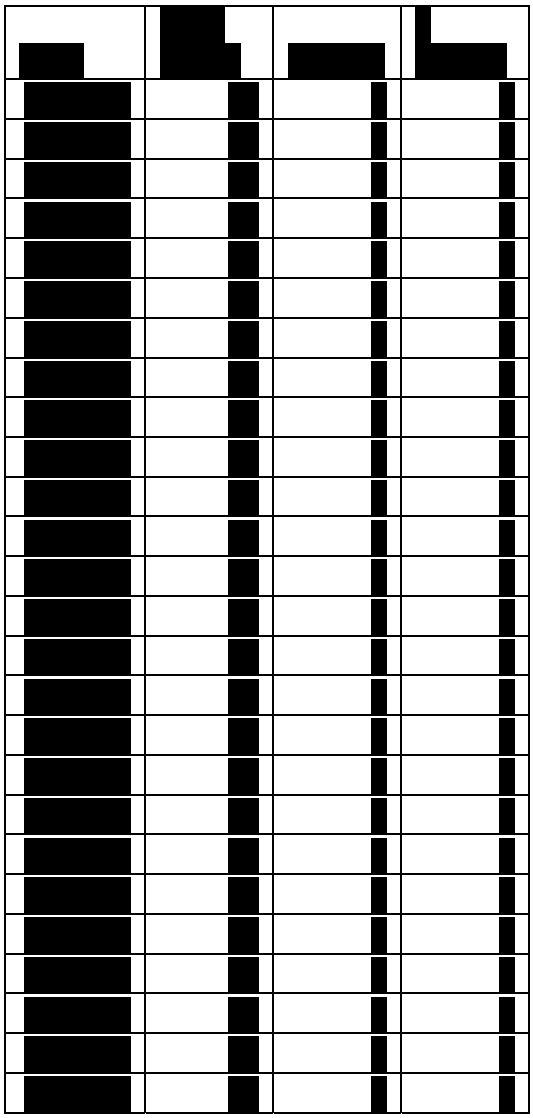


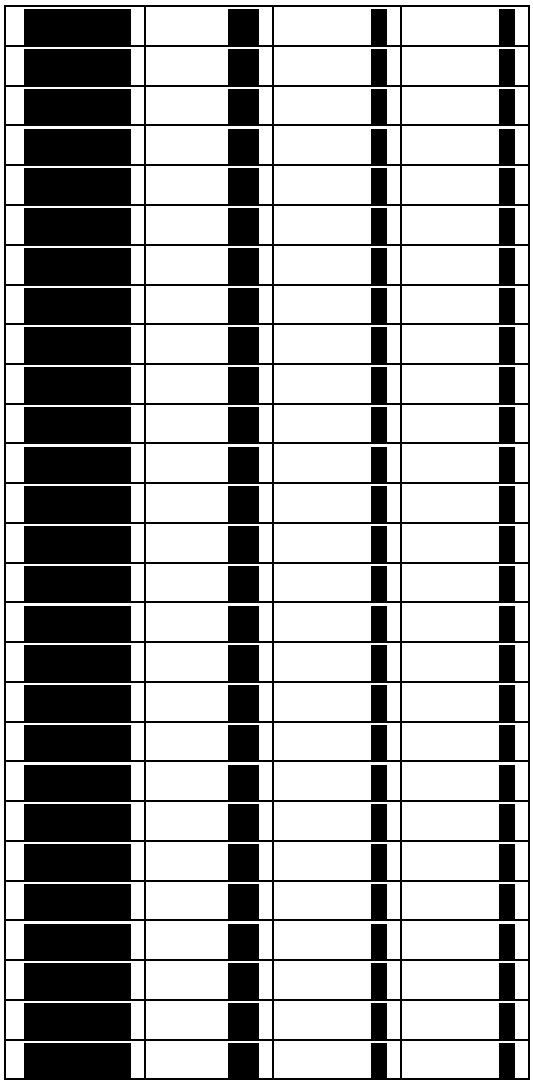
CABOSUN cabozantinib PFS KM data:



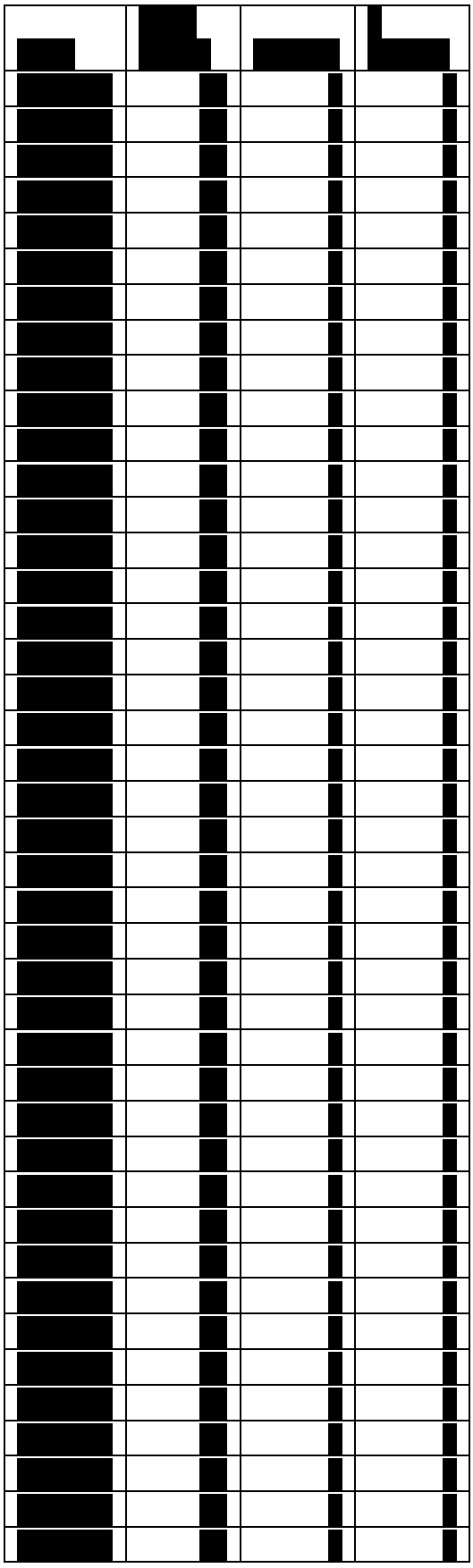


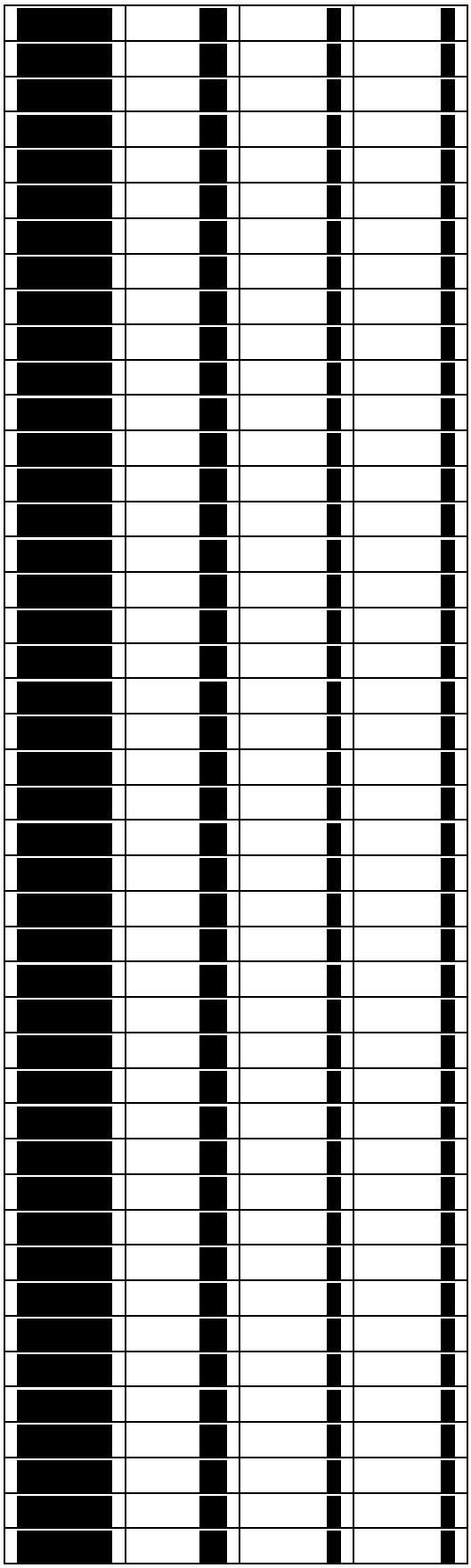
CABOSUN sunitinib PFS KM data:

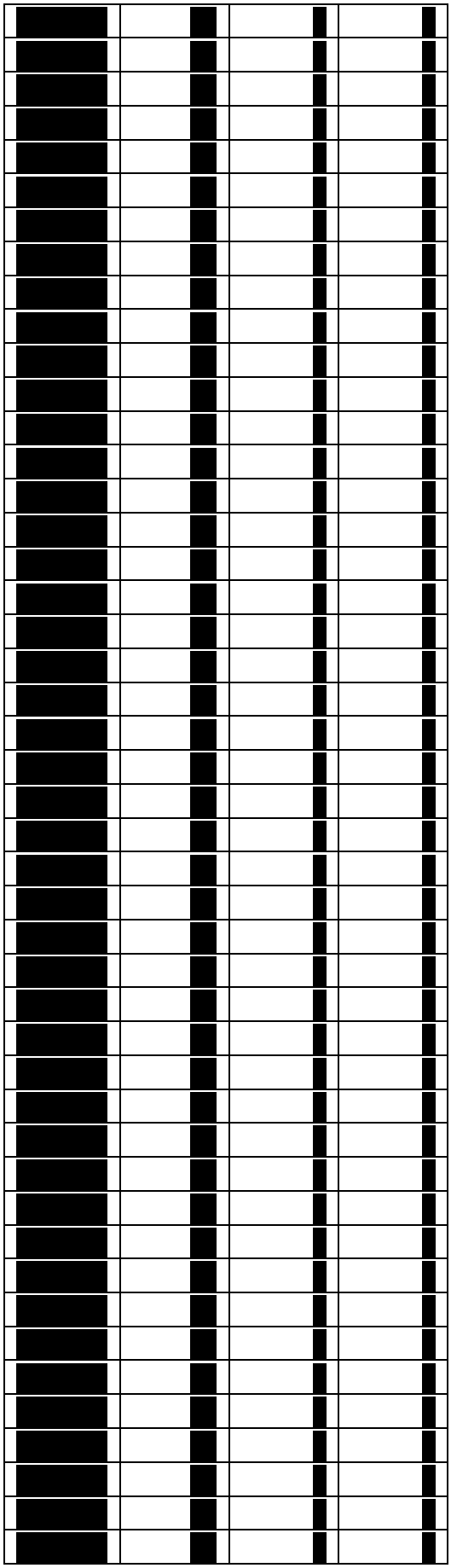


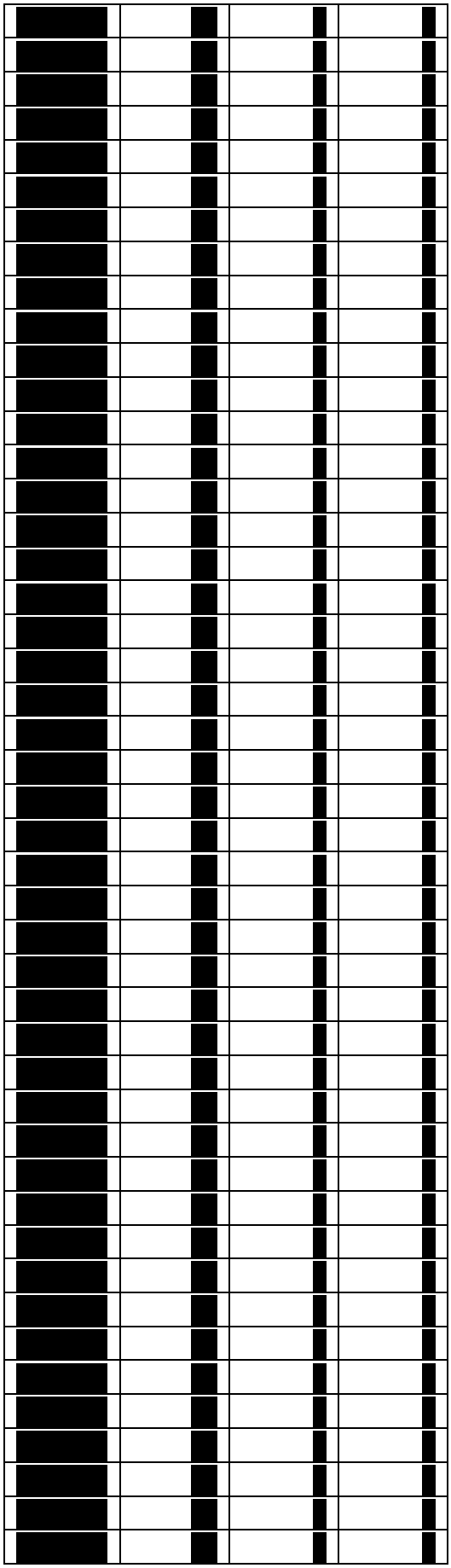


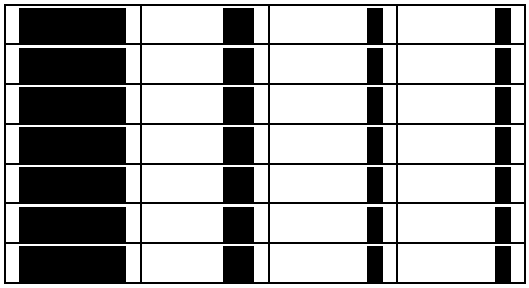
COMPARZ pazopanib OS KM data:

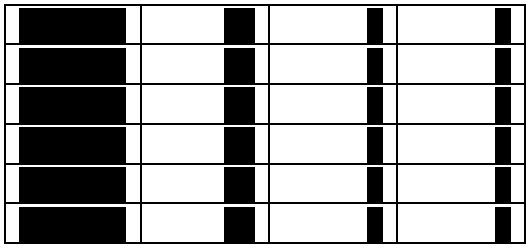




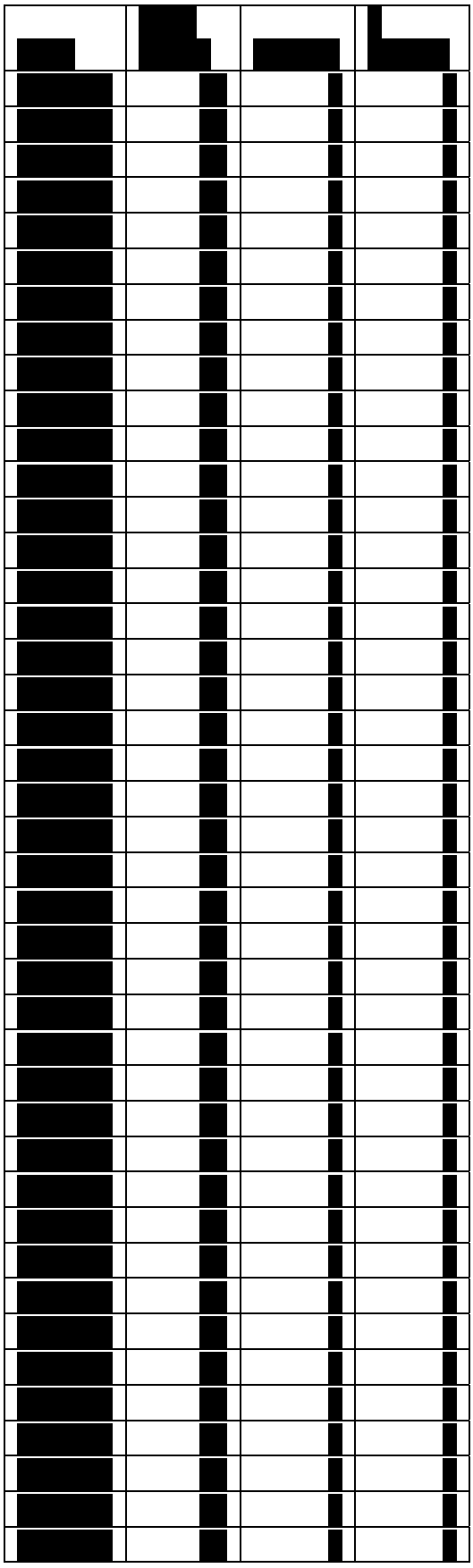


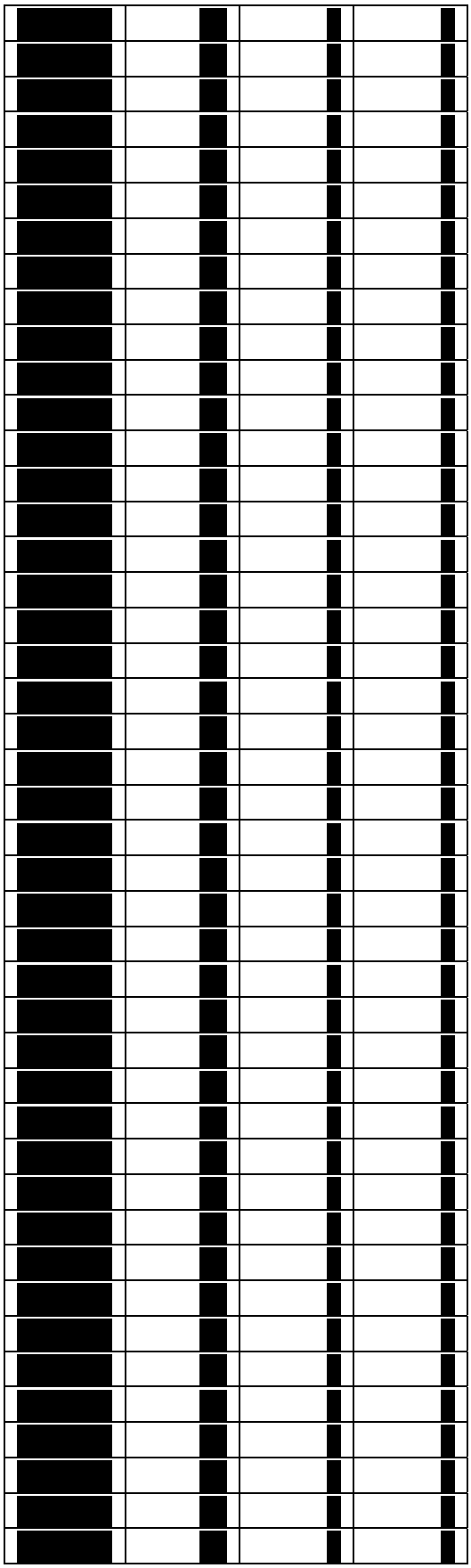


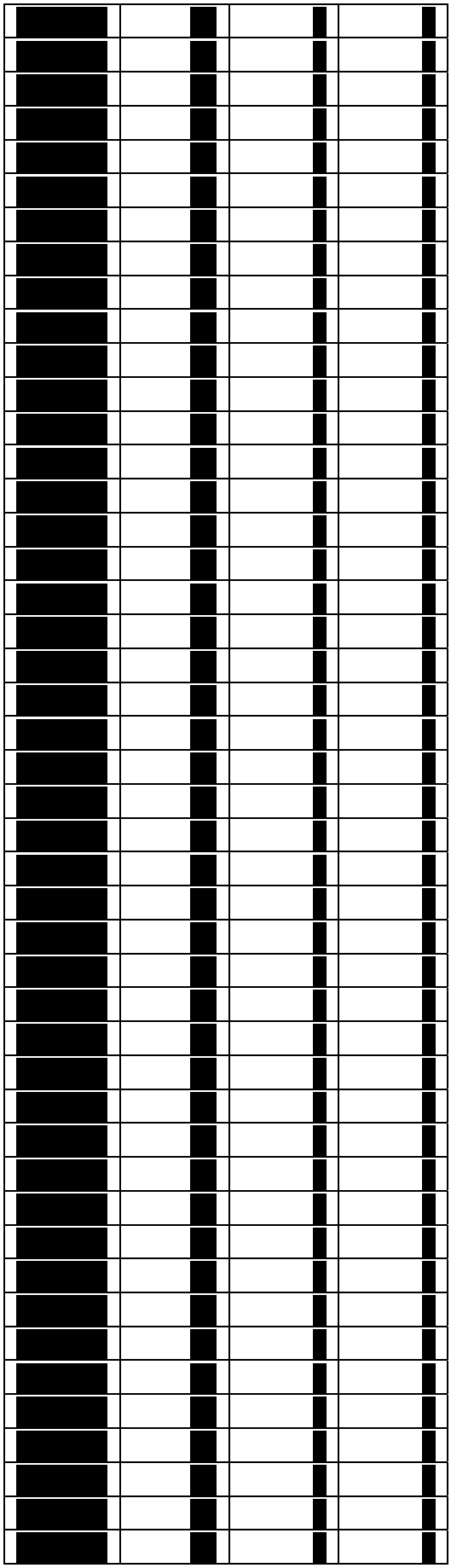


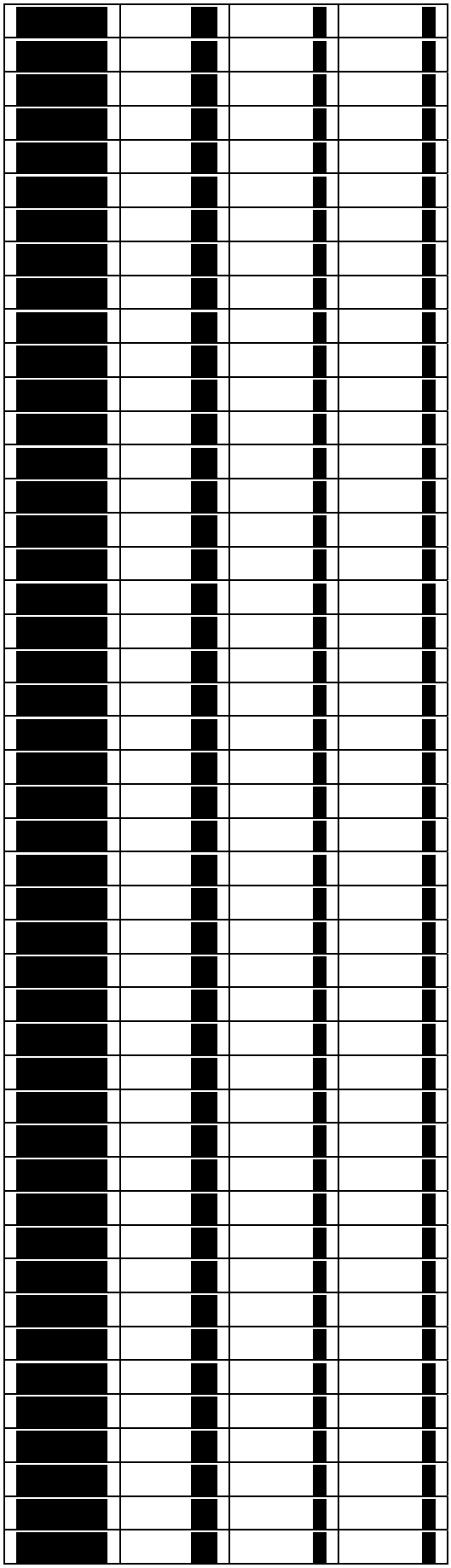


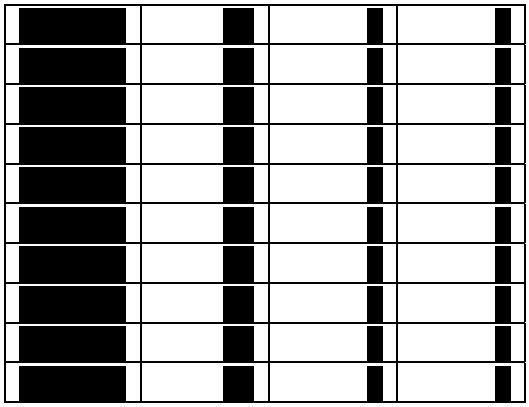
COMPARZ sunitinib OS KM data:

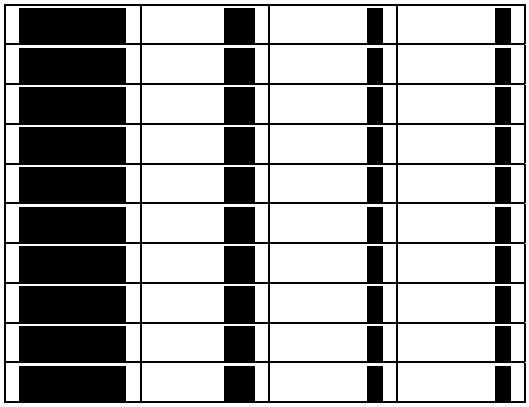




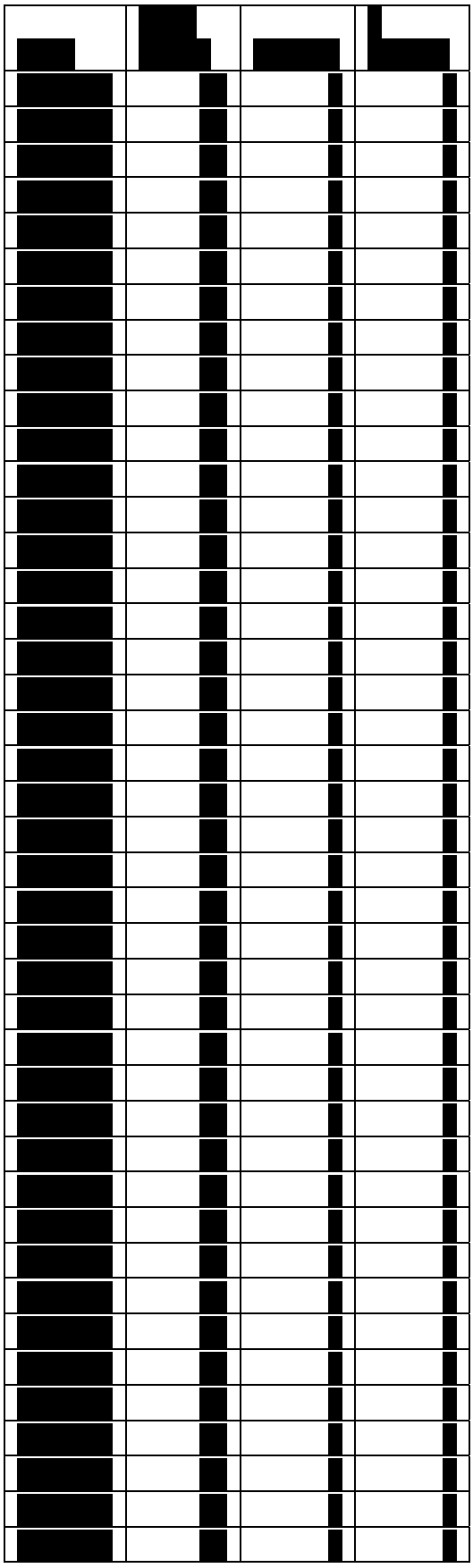


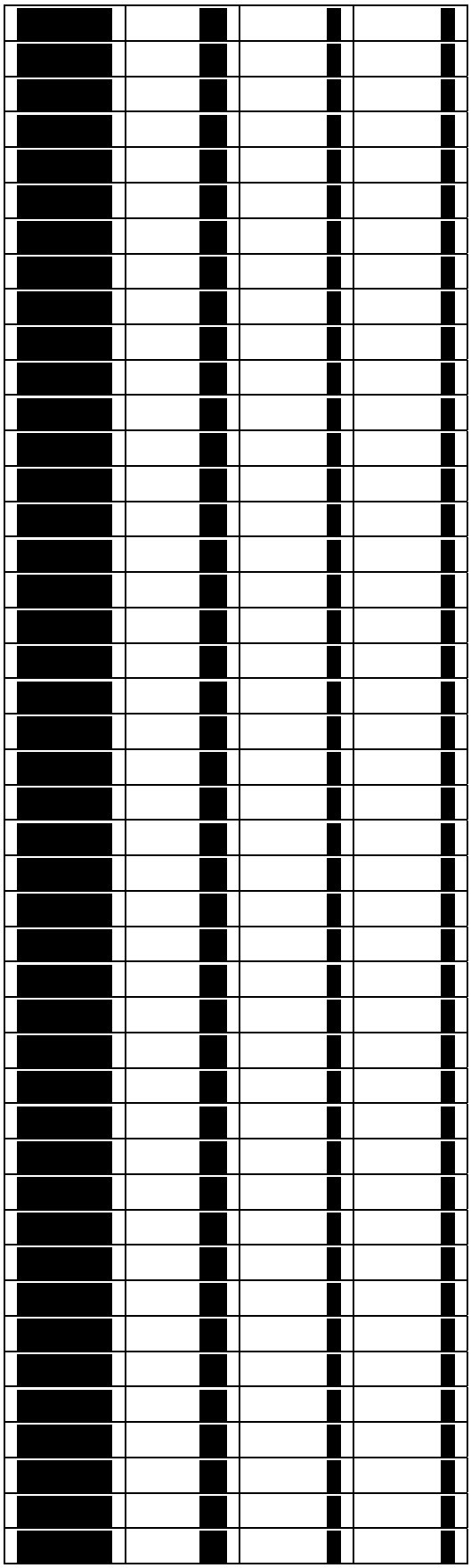


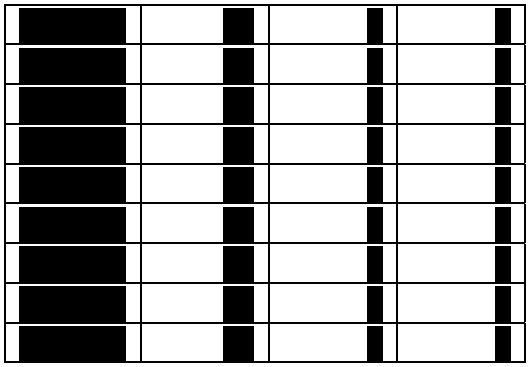


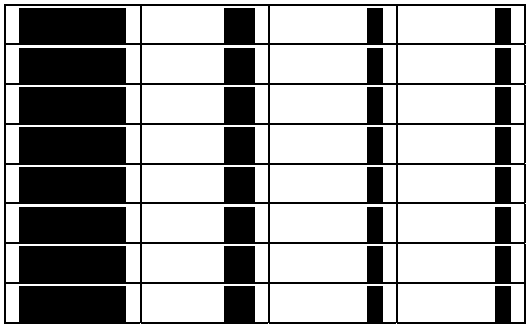


COMPARZ pazopanib PFS KM data:

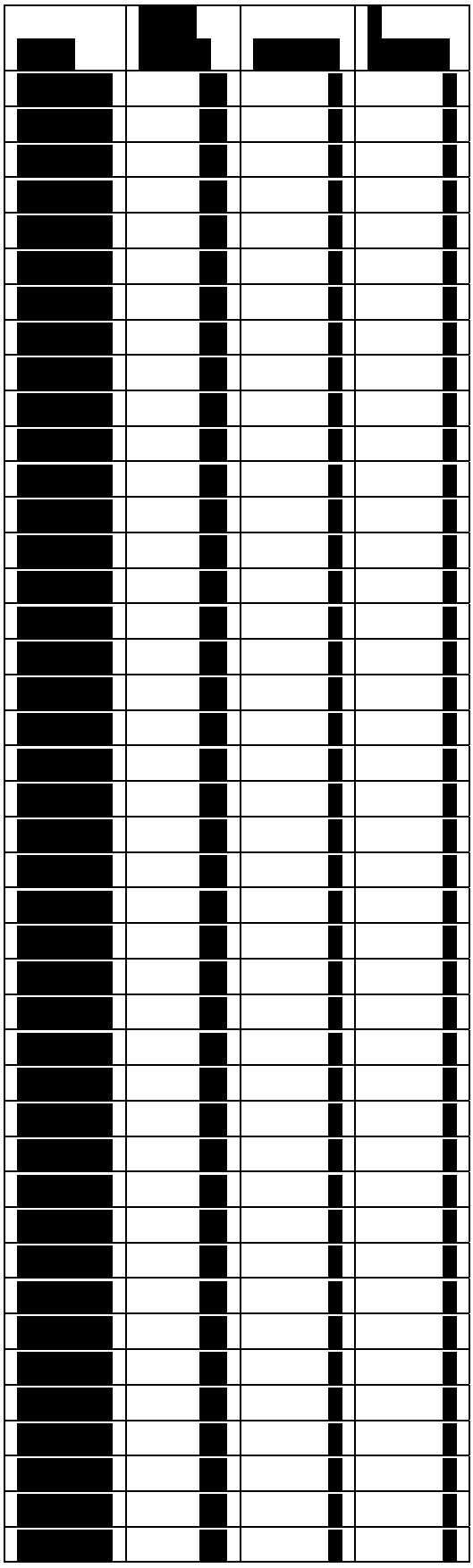


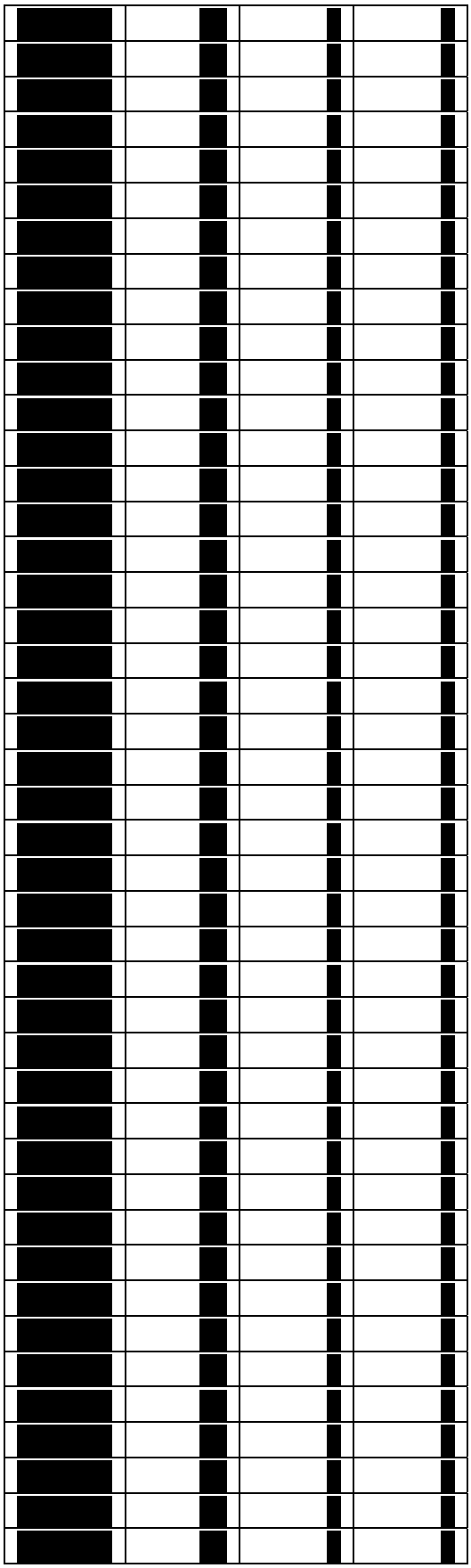


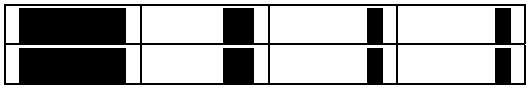


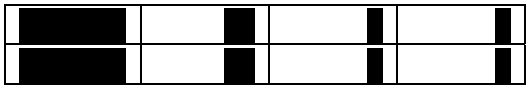


COMPARZ sunitinib PFS KM data:









B2.

Please justify your choice of exponential distribution for the direct comparison of overall survival between cabozantinib and sunitinib in the base case. This is inconsistent with the conclusion that the proportional hazards assumption does not hold for the CABOSUN OS data. The use of separate exponential curves also appears to give a poor fit for the Kaplan–Meier data in Figure 13.

Response: The use of exponential distribution for the direct comparison of overall survival is mainly driven by the clinician's opinion on the long-term survival. Based on the statistical fits of CABOSUN patient-level data analysis, the lognormal distribution is the best fit to cabozantinib and the Gompertz distribution is the best fit to sunitinib. However, the lognormal was not considered as a reasonable assumption for long-term OS extrapolation by the oncologists consulted, as it had a high survival rate at Year 10 (10% for cabozantinib and 8% for sunitinib). The oncologists consulted considered that the lognormal distribution gave too optimistic results for overall survival and the exponential distribution was more reasonable in our model.

We concede that the exponential distribution is not the curve with the best fit to the Kaplan-Meier data. However, the extrapolation beyond the clinical trial duration played a key role in the choice of base case distribution, and for this reason we chose the exponential distribution.

Also, it is not recommended by NICE DSU Technical Support Document 14[5] that the intervention and its comparator are modelled using different types of distributions. Therefore, the exponential distribution was applied for cabozantinib and sunitinib as the base case instead of e.g. Gompertz.

B3.

Priority question. Please can you clarify your assumptions regarding time to treatment discontinuation for pazopanib in the model. On page 142 (B.3.11) it is stated that; "Time on treatment for pazopanib was not identified in the published literature. For pazopanib, the median treatment duration and median PFS data of pazopanib were used to estimate the TTD for pazopanib". We assume this is a typographical error and that the median treatment duration and PFS data of sunitinib were used as an estimate, please can you confirm this assumption? This assumption does not seem to match the formulae in the model (column S in 'E.Pazo.ITC' and 'E.Pazo.FP).

Response: We confirm that this is a typographical error in the submission document. But the median treatment duration and PFS data of sunitinib were not used as an estimate either. Instead, the TTD of pazopanib was assumed to be the same as the one of sunitinib, because the COMPARZ study reports the same mean treatment

[5] Latimer N. NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. NICE Decision Support Unit. 2013. http://www.nicedsu.org.uk. Accessed 8/24/2016.

duration for both sunitinib and pazopanib. The formulae in Column S of 'E.Pazo.ITC' and 'E.Pazo.FP' are consistent with this assumption.

B4.

The utility values reported in Table 43 of the company submission for progression-free (0.78) and progressed (0.705) states are different from those used in the model: 0.77 for progression-free and 0.72 for progressed. Please clarify these differences.

Response: The correct utility values are 0.78 and 0.705. We will correct this issue in the model and provide an updated version.

Section C: Textual clarifications and additional points

C1. Table 15: Please clarify the meaning of footnote d, which appears ambiguous.

Response: Footnote d 'Progressive disease as best overall response' refers to the proportion of patients whose best overall response to treatment with regard to tumour response was classified as 'progressive disease'.

C2. Table 13 footnote states "HR <1 indicates OS in favour of sunitinib". We presume this should say "HR >1" or "… in favour of cabozanitinib"?

Response: The footnote should state 'in favour of cabozantinib'.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma [ID1208]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	

2. Name of organisation	Kidney Cancer Support Network
3. Job title or position	Head of Medical Relations
4a. Brief description of the organisation (including who funds it). How many members does it have?	Kidney Cancer Support Network (KCSN) was founded in 2006 by cancer patients/survivors Rose Woodward and Julia Black, who started by providing practical and bespoke support to individual patients for access to life-extending cancer drugs to treat metastatic kidney cancer. Empowering patients to take an active role in their own health care, and, more generally, in decisions affecting the choice, provision and quality of cancer services throughout the UK, remains the top priority for KCSN. Over the years, KCSN has grown considerably, with a membership of over 1000 kidney cancer patients and carers, and a further 800+ active and committed patients and carers on its confidential social networking sites. KCSN is unique; until recently it operated as a voluntary organisation, totally patient-led and managed by the patients and carers it represents. Although KCSN remains patient-led, the group is now a registered charity, which enables it raise the funds to better meet the growing needs of the UK kidney cancer community. KCSN is funded by grants from trusts/foundations/grant-making organisations and the pharmaceutical industry, in addition to donation from patients and fundraising events/activities carried out by the kidney cancer community in the UK.
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and	When gathering the information for this submission, we specifically asked for patient and carer experience of using cabozantinib through our closed social media channels. Over 800 patients and carers use these channels to communicate on a regular basis, and we receive in the order of 500 posts a day on our closed Facebook group.

carers to include in your submission?	
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Kidney Cancer Support Network (KCSN) is a patient-led kidney cancer charity with the largest and most active patient and carer membership across the UK. As such, we feel we are in the strongest position to feedback how metastatic renal cell carcinoma (mRCC) affects the day-to-day lives of people living with this disease. In 2014, there were more than 12,500 new cases of kidney cancer diagnosed in the UK (34 cases diagnosed every day) and kidney cancer is the seventh most common cancer affecting British people (2014). Kidney cancer accounts for 3% of all new UK cancer cases (2014). In 2014, nearly 4,500 people died from the disease and about 40% of kidney cancer patients will be diagnosed with late stage disease. In these cases, it is estimated that around only 10% of people will survive for five years or more (Cancer Research UK). It is difficult to remain positive in the face of figures like this. Metastatic RCC is a devastating disease and is currently incurable. The majority of mRCC patients are forced to give up work because of the disease itself, and current treatments are very debilitating. This brings with it enormous financial pressures for the patient and their family (and additional costs to the state) and can precipitate psychological problems; depression, loss of confidence and self-worth. Patients may suffer constant pain from metastatic tumours in the brain, bones, lungs, liver, and other more rare sites. Patients with bone metastases are at risk of bone breaks and spinal cord compression. Metastases in the lungs can lead to breathlessness, and persistent coughing, while spread of the cancer to the brain can lead to severe and debilitating headaches, confusion and, in some cases, paralysis. Kidney function is often compromised and patients find daily living difficult, often needing periods of rest during the day. Current first-line treatments offer an important, but sometimes short-lived period of stability, but not all patients respond to these treatments and most patients become refractory after a period of time. Patients diagnosed with hereditary kidney cancer or rare RCC subtypes (non-clear cell RCC) currently have very limited treatment options.

	Biomarkers for the treatment of RCC are yet to be identified, and unfortunately clinicians are not able to predict which patients will respond to which drug. Therefore, selection of the most effective treatment for individual patients is accomplished by trial and error. Clinicians in the UK should have the ability to choose the most effective treatments for individual patients from those available. Without a choice of treatment alternatives, most patients will face disease progression, including worsening of symptoms, such as severe pain, fatigue and shortness-of-breath. Patients require choice in therapy to continue managing their disease, and to maintain quality of life. Patients tell us that psychological support is very difficult to access, and many patients are prescribed anti-depressant drugs to help manage their mental as well as physical clinical situation. Sexual function is affected for both male and female patients, and family life suffers as a result. Kidney cancer cases are rising year-on-year and there is a need for first-line treatment with better overall survival rates than currently exist, especially for difficult-to-treat rare subtypes of RCC. The impact of a terminal diagnosis on the family, as well as the patient, also needs consideration; these families need support during the most difficult time in their lives when a loved one is diagnosed with a terminal disease.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	The current treatment pathway for mRCC is surgery (either radical or partial nephrectomy), followed by either sunitinib or pazopanib in the first-line setting, and axitinib, everolimus, cabozantinib or lenvatinib plus everolimus in the second-line setting, all of which are oral medicines and have similar modes of action (tyrosine kinase inhibitors (TKIs) or mTOR inhibitors that block angiogenesis). Nivolumab is also recommended for use within NHS England for second- or third-line treatment of mRCC. Nivolumab is an immunotherapy (anti-PD-1), which is administered as a biweekly intravenous infusion, requiring outpatient hospital treatment (chemotherapy chair resources), and the associated travel time and expense for the patient and carer. We have extracted the following details from statements submitted to the KCSN by patients living with mRCC. Using currently available drugs, many patients suffer with: • Extreme fatigue

	• Severe hand and foot syndrome which can leave patients unable to walk • Intestinal problems (chronic diarrhoea) • Pneumonitis requiring hospital treatment and cessation of treatment • Severe mouth ulcers causing problems eating and drinking • Nausea and vomiting, which can also cause problems taking the medication • High blood pressure (hypertension) • Hyperthyroidism All the above side effects require additional medicines to help patients manage the drugs and/or tumour pain, which require opioid prescriptions respectively. Costs for additional medicines to mitigate the side effects of these targeted therapies should be taken into account. Other less serious side effects, which still affect the patient's quality of life, are loss of taste, hair loss and change of hair colour, depression, loss of libido, and inability to drive. In some cases, treatment can affect a patient's quality of life to such an extent that clinicians recommend a dose reduction, and some patients are even advised to stop treatment as a result of severe side effects. Patients are aware that these treatments are life-extending drugs, but they continue to look for drugs with different modes of action, which can give improved overall survival with better quality of life. For a lot of patients, the most important treatment outcome would be no evidence of disease, i.e., a potential cure for their kidney cancer. The hope of achieving this outcome spurs patients on to continue to take current medication, despite significant toxicity, and to search for alternative, more effective treatments that can extend overall survival. Failing to achieve no evidence of disease, tumour shrinkage or disease stability would be the next best outcome for patients. In addition to treatment outcomes, quality of life is also an important consideration for many patients. Most patients would prefer a treatment that allows them to continue to lead as normal a life as possible, and to contribute both socially and economically to their communities: *"The extra years, which the drugs give me, enable me to carry on working, using the accumulated knowledge and experience, gathered through my working life, for the benefit of the various ……. enterprises which I manage……..I'm making a hugely positive contribution to society, and the wider*

	economy, and I wish to be able to carry on with this and more importantly to ensure that others, whatever their circumstances, will have the same opportunities". *"……has enabled me to enjoy every day, do 3 or 4 days voluntary work a week and to care for my elderly parents. The side effects for me have been milder than many people but the fear of diarrhoea striking all through the day makes travelling and working very difficult. I would like a treatment without digestive effects, little fatigue and control of growths……".* Although less serious than some of the side effects to current first-line treatment, some patients find the changes to their appearance caused by these treatments distressing: white, thinning hair, and pale skin make them feel nearer to death and also singles people out as cancer patients. Some of the current first-line treatments can also cause issues with the thyroid gland, blood pressure, and cholesterol levels. From a psychological point of view, knowing that you have stage 4 cancer and that there are possibly more effective treatments that you are not able to access is very difficult for patients. Carers seem to find this even harder, as they live with a guilt of not being able to do everything they can for their loved one. Access to a choice of treatments in the first-line would enable patients and their families to know they had tried their best to beat the cancer, leading to better family relationships and a subsequent improvement in quality of life and wellbeing for the patient. Nowadays, kidney cancer patients do not exist in silos. They communicate widely within online patient communities; international discussion forums exist where patients talk to one another daily, and patients are more aware of the experiences of others, including their access to innovative treatments, quality of life, and treatment successes and failures. News about lack of access to effective medicines ripples out to other patients and families, destroying their hope and positivity. Information about new treatments is readily available to patients around the world on the internet. Patients and clinicians are right to expect NICE and the pharmaceutical industry to find a way to bring new and innovative treatments to kidney cancer patients in England to improve outcomes, so that patients in England have the same choices as patients in other countries.
8. Is there an unmet need for patients with this condition?	Current first-line treatments are not effective at treating bone metastases, which are often treated with radiotherapy. There is a significant unmet need for a safe and effective treatment for RCC patients with spread to their bones.

	There is an unmet need for a first-line treatment that improves overall survival and allows patients to live a good quality of life without the incumbent debilitating side effects of current first-line treatments. There is also a significant unmet need for effective and safe treatments for people with hereditary kidney cancer or rare RCC subtypes (non-clear cell RCC), who currently have very limited treatment options.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Cabozantinib has been proven to be a clinically effective drug, and designated a 'promising innovative medicine' for advanced RCC by the Medicines and Healthcare products Regulatory Agency (MHRA) in 2016. Also, cabozantinib was designated a breakthrough therapy by the FDA for the treatment of advanced RCC in 2015. Cabozantinib's efficacy in the first-line setting was proven in the CABOSUN study, where cabozantinib was compared with standard first-line treatment with sunitinib. First-line treatment with cabozantinib reduced the risk of disease progression or death by 52% compared with sunitinib for patients with advanced RCC. Median progression-free survival was 8.6 months with cabozantinib versus 5.3 months for sunitinib. Cabozantinib is the first tyrosine kinase inhibitor to act on multiple tyrosine kinase receptors, including c-MET, VEGF2, AXL and RET. Its c-MET activity may explain its effectiveness against bone metastases, since MET appears to be an important growth factor in the bone microenvironment. The following statement from the husband of a patient highlights the importance to patients of cabozantinib's efficacy against bone metastases: *"…..CT and MRI results …… yesterday gave excellent news confirming her 10-off* [sic] *spinal bone Mets being reported stable. This is a great result having halted the disease given she only recently commenced her Cabozantinib treatment on 23/11/16; at a time when the bone progression appeared aggressive, i.e. with 3 lytic bone Mets being reported by CT scan on 21/10/16 increasing to 10 Mets reported from an MRI scan on 19/12/16.* *"……. the immediate issue was rapidly developing bone mets (i.e. crocodiles nearest the boat, so to speak). Since Cabo was the only 'available' agent that has a pathway able to specially target*

	bone Mets, then this became OUR first choice …... Note: we had overturned the originally advised preference ranking order for Axitinib, Nivolumab and lastly Cabozantinib." It seems that cabozantinib may be particularly effective for treating patients with bone metastases: however, further clinical evidence is needed to confirm these anecdotal findings. The following statements are from a patient and the wife of a patient with a bone metastasis, and demonstrates how well informed advanced RCC patients are and how important access to cabozantinib in the first-line is for them: *"Three years after a nephrectomy for RCC, I became aware of bone pain in my femur, which subsequently broke due to a single site metastasis that had become so large there was very little bone remaining. Following surgery, in December 2014 I was started on Sunitinib. At that time I had no other mets, and that is still the case, so Sunitinib has been successful in preventing spread, however, it has had no measurable impact in reducing the bone met, over 2 years later. Sunitinib, like the other currently approved drugs is not greatly effective on bone mets. However Cabozantinib has clear data demonstrating that it can be highly effective in shrinking and removing altogether bone metastases. For me, that could mean achieving NED, which result in a big saving in no requiring further expensive treatment* [sic]. *"This is the only drug currently available that is so effective on bone mets and therefore for patients like myself it is essential that this drug is approved for use at least in the second line setting to offer real hope to patients with bone metastases. I would therefore urge NICE to approve this new drug as soon as possible"* *"My husband has run out of options for surgery on his maxilla area without it compromising his eye. His other secondaries are kept under control and after nearly 7 years he is stable. He needs a drug, which works on bone metastases as none of the current drugs appear to have any measurable success and sadly kidney cancer often goes to hips and spine as well as other areas."*

Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	The main disadvantage of cabozantinib as far as patients and carers are concerned is the side effect profile for cabozantinib. However, clinicians and patients have a number of years of experience dealing with TKI adverse events, and consider the benefits of improved overall survival and effectiveness against bone metastases outweigh the inconvenience of adverse events. The METEOR trial confirms that adverse events with cabozantinib are as expected for a TKI for the treatment of advanced RCC. The proportion of patients reporting an adverse event was similar for cabozantinib and everolimus, the most common adverse events for cabozantinib being diarrhoea, fatigue and nausea. As for other TKIs, such as sunitinib and pazopanib, adverse events are managed by dose reductions or 'drug holidays', and patients are prepared to accept this inconvenience for the benefits the drug provides: *"Just been to see you* [sic] *oncologist, cabos* [sic] *not been very kind sore mouth peeling hands and high blood pressure decided to give me a two week break bit disappointed but it's a new drug I probably know more than they do they follow the drug advice and play it by ear starting again on January 1st on 40 mg down from 60 heres* [sic] *hoping the side effects wear off quickly and I can enjoy my turkey."* *"This is an update on cabozantnib* [sic] *after the first month on 40mg the side effects are better with the reduced dose the worst seems to be the sore mouth and loss of taste although I have a number of cures nothing gets rid of it completely I also got unexplained muscle pain and a need to sleep at least twice a day I have not had full diarrhoea something in between. I am glad to say this is all wearing off for now and I am starting to feel good again so to sum up nothing terrible has happened and I have coped the real test will be when I have a scan and we find out how successful it has been".* The following statement from the husband of a patient taking cabozantinib highlights the proactivity of patients when managing side effects. This patient has been taking cabozantinib for nearly 3 months and now has stable disease: *"As regards side effects, these were getting too tough last week and coincidentally combined with a*

	urinary tract infection, prompting antibiotics and an unscheduled intermission in her Cabo treatment. Planing [sic] *to resume Cabo this Saturday but shall need to keep the proactive drinking and exercise regime to alleviate toxicity effects and fatigue. Diarrhoea has been problem but treated with Imodium and codiene."* *"Just finished first week on cabo and to be honest I cant believe how well it has gone I started to feel better almost immediately and it has carried on the pain in my back has reduced and my mobility improved so far few problems very sore mouth slight blood pressure and diabetes variations but otherwise ok I can't believe it if this carries on it could be the wonder drug were* [sic] *all hoping for. 2nd week on cabo not as good as first side effects started Monday sore mouth is worst got gelclair difflam but there* [sic] *not very effective at the moment and fatigue its messed my sleep patterns up ….. but I suppose side effects must mean its working just trying to keep positive.* *"I have now had the official report on my scan tumour in my lung which was 2cm is now 1cm and although it is difficult to see because of the metalwork on my spine it is stable this is great news. Because bone doesn't regenerate itself stability is something as far as cabozantnib* [sic] *its self it has been challenging sore mouth peeling hands dioreaha* [sic] *and muscle pain and fatigue my oncologist had been to a presentation by the drug company at which they made light of the side effects but she stated that all patients …. taking the drug were finding it challenging however when you find it is working then it is all worthwhile. I am on a 2 week break which may be necessary to tolerate as with sutent and also further reduction to 20mg as you can imagine I am delighted and it is another weapon in the armoury something to give hope to all."* We understand that cabozantinib is expensive, and we appreciate the budgetary implications, but nonetheless NICE and the manufacturers must negotiate and find a way to make this innovative drug available to the patients who need it; failure to do so would be seen as failure of professional competence. NICE and the manufacturer need to think outside the box to agree an alternative funding mechanism, and work collaboratively to negotiate an acceptable patient access scheme to ensure kidney cancer patients who need it can have access to this latest clinically effective drug. *"My dad's consultant has suggested that should nivolumab stop working then this would be the next step. He specifically mentioned that Cabozantinib was more effective on bone mets than other lines of treatment, which we took as a positive since dad has mets on his spine. If this wasn't an option I think we'd be at the end of the line as dad has had IL2, sutent and axitinib prior to*

	nivolumab. It really would be a matter of life and death and to know that there is something there that could extend life but wasn't available would be heart breaking. I know there has to be assessments around cost versus impact, but given dad's history it might have been felt that nivolumab wouldn't work when it has - he's been on it for almost a year now. Some weren't as lucky as dad and missed nivolumab. I'd hate to see this happen again."
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	From the anecdotal evidence we have gathered from the advanced RCC patients in the Kidney Cancer Support Network currently taking cabozantinib either through the Managed Access Programme or as a second-line treatment through NHS England, this drug offers hope and an alternative effective treatment to patients who have spread to their bones: *"I [am] on Cabozantinib but have only had 4 weeks on it (with a week's gap in between as I had some radiotherapy)…….. the doctor I saw noticed that a visible lump on my jaw (on the muscle) which I had been told, after a biopsy, was down to the kidney cancer, had gone down quite a bit. Obviously this doesn't mean much overall but he did say that they've found when the Cabozantinib works, it works quite quickly so it looks like a positive sign at least…. So far the side-effects have been quite manageable but ……. I don't know how it will be after a few weeks without a break. …… I have some bone mets…… I previously had 18 months on Everolimus, which suddenly stopped working, then a few weeks on Nivolumab, which the doctors felt was enough to show it wouldn't work for me so I'm not sure what the alternative would have been."*

Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	None
Other issues	
13. Are there any other issues that you would like the committee to consider?	Cabozantinib is already available as a first-line treatment for advanced RCC in the USA. Currently, UK cancer survival rates trail about 10 years behind other comparable European countries, including Italy and Austria. If the UK is to improve patient outcomes, including the patient experience as well as overall survival, it is vital that innovative new drugs are made available to patients in order that they have the best possible care. If these drugs are not made available, it leaves UK patients at a major disadvantage in terms of the availability of innovative cancer treatments; these patients are likely to die prematurely compared to kidney cancer patients in other rich countries. A contributory factor to poor survival rates in the UK is possibly due to the restrictions in clinical choice brought about by UK regulatory authorities. In the absence of biomarkers for the treatment of RCC, clinicians are not able to predict which patients will respond to which drug, and drug selection is accomplished by trial and error. Clinicians should have the ability to choose the most effective treatments for individual patients from those available, and without cabozantinib, the clinician's choice of treatment is seriously compromised. Without treatment alternatives in the first-line most patients will face disease progression. A choice of treatment is paramount for the effective management of the progression of this disease and maintenance of quality of life: *"Whilst I have not had direct experience of taking Cabozantinib as I am still responding to Pazopanib, I have read both the clinical trial reports and real world patient experience. I believe that this would form a useful addition to the portfolio of drugs available to clinicians and will be*

	especially useful for those patients with bone metastasis. The addition of more potential drugs would introduce more competitive pricing between suppliers." Current first-line treatment options are not effective for everyone. Undue restrictions in accessing cabozantinib would simply add unnecessary additional burden to patients with a terminal diagnosis. Having more choice in the first-line setting would enable patients and oncologists to individualise treatment plans according to specific disease/treatment history and contraindications, thereby enabling the best possible quality of life for the patient. Cabozantinib will also address the massive unmet need for an effective treatment option for bone metastases in the first-line setting. The following statements are from a patient carer and two patients talking about the importance of having a choice of treatment: *"Another important consideration to factor is that some drugs can work better later in the cycle of this disease, what I mean is - research supports Nivo tends to be more effective when your cancer is more mutated, so had we chose* [sic] *this option now (possibly too early) then it may have not worked and we could have lost a valuable treatment option needed later."* *"I have used sutent, pazopanib and now axitinib for almost five years. When Axitinib is done, I want to be able to turn to Cabozantinib as I have a bone met. Please give me the choice."* *"In response to cazantinib* [sic] *not being approved by NICE, this is a drug that had been mentioned to me as a next step to help keep my kidney cancer at bay, it could give me valuable extra time with my two young daughters aged 4 & 2 years old. Without this medication my girls could lose their mummy too soon & they don't deserve that. This could help so many people live longer, everybody is worthy of that chance. Please think again."* Choice of treatment is also important when it comes to drug combinations. Cabozantinib is already being tested in combination with immunotherapy, and could prove to be a formidable treatment regimen, if successful. The following are some thoughts of a patient carer on the subject of drug combinations: *"…….. it could open opportunities to be within the mix of combination, e.g. it could be beneficially combined with Nivolumab and Iplimumab* [sic] *to improve the chance of a complete response. I believe this specific combination is being trialled in USA now."*

14. To be added by technical team at scope sign off. Note that topic-specific questions will be added only if the treatment pathway or likely use of the technology remains uncertain after scoping consultation, for example if there were differences in opinion; this is not expected to be required for every appraisal.] if there are none delete highlighted rows and renumber below	

Key messages

15. In up to 5 bullet points, please summarise the key messages of your submission:

- Cabozantinib is the first inhibitor acting on multiple tyrosine kinase receptors to show efficacy in metastatic RCC, and has been designated a 'promising innovative medicine' for advanced RCC by the MHRA in 2016 and a breakthrough therapy by the FDA for the treatment of advanced RCC in 2015. Cabozantinib is available as a first-line treatment for advanced RCC in the USA
- Cabozantinib could be used to address an area of significant unmet need for a safe and effective first-line treatment for bone metastases
- Cabozantinib is a clinically effective treatment for advanced RCC in the first-line setting, and has proven to be more effective at extending progression-free survival compared to standard first-line treatment with sunitinib

- Adding the cabozantinib as a choice in the first-line enables patients and clinicians to individualise treatment plans to better control this disease and maintain a high quality of life
- Cabozantinib could be used to address an area of significant unmet need in the treatment of non-clear cell RCC

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Professional organisation submission

Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma [ID1208]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**NCRI-ACP-RCP**

3. Job title or position	RCP registrar
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or	Palliation of metastatic renal cell carcinoma

disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Progression Free Survival benefit > 6/12
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	1st line: Pazopanib or Sunitinib
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	ESMO guidance

• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Well defined due to NICE appraisal decisions defining access to treatment
• What impact would the technology have on the current pathway of care?	It would provide a potentially improved option for the first line treatment of metastatic RCC for those patients requiring tyrosine kinase inhibitor (TKI) based treatment
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Currently used 2nd and subsequent lines of treatment. This would allow 1st line use.
• How does healthcare resource use differ between the technology and current care?	No significant difference
• In what clinical setting should the technology be used? (For example, primary or secondary	In specialist oncology clinics

care, specialist clinics.)	
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	The CABOSUN study demonstrated superiority of cabozantinib vs sunitininb in the 1st line setting. This was however a small exploratory study and therefore the magnitude of benefit has some uncertainty.
• Do you expect the technology to increase length of life more than current care?	Unknown
• Do you expect the technology to increase health-related quality of life more than current care?	No
12. Are there any groups of people for whom the technology would be more or	There are no predictive biomarkers available

less effective (or appropriate) than the general population?	
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	No
14. Will any rules (informal or formal) be used to start or stop treatment with the technology?	Treatment until progression usually measured on CT scans or treatment limiting toxicity

Do these include any additional testing?	
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	The Quality-adjusted life year (QALY) is insensitive to changes in quality of life in oncology patients. The Utility score is particularly insensitive. The committee are aware of this.
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	No
• Is the technology a 'step-change' in the management of the	No

condition?	
• Does the use of the technology address any particular unmet need of the patient population?	Yes – we require more effective disease control
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Cabozantinib has a toxicity profile that is similar to other TKIs. Only oncology specialists with appropriate experience should manage patients on the drug.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they	CABOSUN study randomised 157 treatment naïve patients to cabozantinib or sunitinib (standard of care). PFS 8.2 vs 5.6 months – HR 0.66. Response rate 33% vs 12&

measured in the trials?	Progression Free Survival – yes Response rate – yes Overall Survival - no
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	n/a
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
20. Are you aware of any new evidence for the comparator treatment(s) since the	No

publication of NICE technology appraisal guidance [TAXXX]? [delete if there is no NICE guidance for the comparator(s) and renumber subsequent sections]	
21. How do data on real-world experience compare with the trial data?	Broadly equivalent, although performance of both arms in CABOSUN was a little inferior to what our experts would have expected.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	

Topic-specific questions	
23 [To be added by technical team at scope sign off. Note that topic-specific questions will be added only if the treatment pathway or likely use of the technology remains uncertain after scoping consultation, for example if there were differences in opinion; this is not expected to be required for every appraisal.] **if there are none delete highlighted rows and renumber below**	
Key messages	

24. In up to 5 bullet points, please summarise the key messages of your submission.

- Cabozantinib is an active drug. We use a lot in 2nd and subsequent lines of therapy
- CABOSUN study shows superiority of cabozantinb vs a current 1st line standard of care – sunitinib.
- CABOSUN is a small clinical trial although was randomised and well conducted. The performance of both standard and experimental arms was a little poorer than that we would have expected.
- First line cabozantinib would be a good option for patients
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NHS England submission on the NICE appraisal of cabozantinib as 1st line therapy in locally advanced/metastatic renal cell adenocarcinoma

1. If the marketing authorisation for cabozantinib is as expected and licenses cabozantinib for the 1st line treatment of patients with intermediate and poor risk locally advanced/metastatic renal cell adenocarcinoma and NICE recommends this option to the NHS, it will only be commissioned by NHS England as 1st line therapy in these patients groups ie it will not be commissioned as 1st line treatment of patients with good risk disease. It would of course remain as a 2nd line treatment option for all risk groups if not previously treated with cabozantinib.
2. NHS England notes that the combination of nivolumab and ipilimumab is also being appraised by NICE at the same time for the same population of patients (intermediate and poor risk patients) at the same point in the treatment pathway (in treatment naïve patients). If both cabozantinib and nivolumab/ipilimumab options are recommended by NICE, then for fit patients who are willing to have intravenous therapies, it is likely that the combination of nivolumab and ipilimumab will be preferred by clinicians and patients. For those who prefer oral treatment and fewer hospital visits, cabozantinib will have a place in therapy. For less fit patients (eg ECOG performance status 2), there will be concern as to the toxicities of both cabozantinib and the nivolumab/ipilimumab combination.
3. NHS England notes the small size of the cabozantinib vs sunitinib randomised phase II trial and the uncertainty that this creates. Also noted are the inferior median PFS and OS durations (5.3 months and 21.2 months, respectively) for sunitinib in this trial versus that in the sunitinib arm (8.4 months and 26 months) of the much larger nivolumab plus ipilimumab phase III study against sunitinib. This is despite similar ratios of intermediate to poor risk patients (4.3 in the cabozantinib trial, 3.7 in the nivolumab/ipililumab trial).
4. The use of nivolumab as a subsequent treatment in both arms of the cabozantinib vs sunitinib study is much lower at 15% than that seen in NHS England.
5. NHS England notes that 1st line TKI in practice in England is currently about 60% (and rising) for pazopanib and 40% (and falling) for sunitinib. Both drugs are considered to be equally efficacious but pazopanib has fewer side effects. It is likely therefore that pazopanib will be better tolerated than cabozantinib as 1st line TKI.
6. If NICE recommends cabozantinib as 1st line treatment, its use in this place in the treatment pathway will largely depend on whether NICE recommends the combination of nivolumab and ipilimumab. If the 1st line immunotherapy is recommended, the use of cabozantinib is likely to be relatively modest; if not, then cabozantinib is likely to find a place in 1st line treatment of intermediate/poor risk patients in the fitter patients.
7. NHS England notes that no quality of life data was collected in the cabozantinib vs sunitinib trial. NHS England regrets that in 2018 a 1st line palliative therapy can

obtain a potential place in the commissioned treatment pathway for renal cancer without direct evidence of the impact on quality of life of patients.

8. NHS England notes that the median survival duration for patients with intermediate risk metastatic disease is rising. In the International Metastatic Renal Carcinoma Database Consortium 2011 publication of patients in the intermediate risk group who were not treated with TKIs (1975-2002), the median survival was 11.5 months. When the IMRCDC model was tested in a population-based study of patients treated with TKIs in the period 2004-2010, the median survival duration of intermediate risk patients was 22.5 months (Lancet Oncol 2013; 14: 141-8). There are several new NICE-approved treatment options which have been brought into clinical practice since 2010. Hence NHS England considers that the median survival duration of the intermediate risk metastatic group will be at least 2 years and the mean survival duration substantially more than 2 years in this group. NHS England also notes that the company-modelled mean survival of the intermediate and poor risk group in the sunitinib control arm of Cabosun study was in excess of 2 years.

[redacted]

[redacted]

May 2018

NICE National Institute for Health and Care Excellence

Patient expert statement

Cabozantinib for untreated metastatic renal cell carcinoma [ID1208]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]
2. Are you (please tick all that apply):	☐ a patient with the condition? ☐ a carer of a patient with the condition?

	☐ x a patient organisation employee or volunteer? ☐ other (please specify):
3. Name of your nominating organisation	Kidney cancer UK
4. Did your nominating organisation submit a submission?	**X**☐ yes, they did ☐ no, they didn't ☐ I don't know
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	**X**☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	**X**☐ yes
7. How did you gather the information included in your statement? (please tick all that apply)	☐ I have personal experience of the condition ☐ I have personal experience of the technology being appraised ☐ I have other relevant personal experience. Please specify what other experience: ☐ **X** I am drawing on others' experiences. Please specify how this information was gathered:
Living with the condition	
8. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Different people will react to living with kidney cancer differently and the challenges they face greatly depend on the stage of their disease. Most people with kidney cancer will receive surgery at some point, which will require a period of recovery. There will be times when the patient and family/carers will be worried about the future and require information and guidance. Waiting for news, scans and procedures can be emotionally draining. Knowledge that there are a variety of treatment options available to them will give them some comfort. Dealing with side effects of drugs can be equally exhausting as the symptoms of the cancer, so finding the balance of treatment and quality of life that is right for each patient is important.

Current treatment of the condition in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS?	The treatment and outcome are very much dependant on how early the kidney cancer has been caught. Ideally the tumour is of an early stage and is removed by surgery or cryotherapy and the patient enjoys a life after cancer. This would always be the preferred treatment. However, if the tumour has spread patients will rely on targeted therapies and immotherapy treatments. Current drug treatments for kidney cancer are very limited in number and have plenty of side effects. Kidney Cancer UK feel that there are significant improvements that could be made in this area. A wider range of options with improved efficacy and fewer side effects. The most commonly used Tyrosine kinase inhibitors (sunitinib and pazopanib) act to extend life and in some cases they work very well and extend life for many years. For others, the extension of life is a matter of months. However, those months can be invaluable for individuals and their families. The recent introduction of nivolumab (immunotherapy) as a NICE recommended 2nd line drug is very good news. Patients have reported back on how effective this drug has been for them especially their quality of life. I think that having combinations of treatments may give alternate options and even better results as a first line treatment. Giving alternate options for patients can be invaluable especially in an era where personalised medicine may be introduced.

	Cabozantinib acts in a similar way to the other TKI's but has been shown to target c-MET and VEGFR2 tyrosine kinases so it could be beneficial as an alternative 1st line treatment. A multitude of treatment options is always desirable.
10. Is there an unmet need for patients with this condition?	Yes there is an unmet need for treatment of advanced RCC it would most certainly improve some outcomes in patients surviving kidney cancer and to be free of cancer for the foreseeable future. We understand that most drug treatments aim to extend the lives of people with kidney cancer and viewing kidney cancer as a chronic disease that can be lived with would be a desirable outcome. Tolerable side effects of a treatment are important if kidney cancer is to be viewed as a chronic disease and patients are to have a good quality of life.
Advantages of the technology	
11. What do patients or carers think are the advantages of the technology?	Benefits of a treatment might include its effect on: • the course and/or outcome of the condition • physical symptoms • pain • level of disability • mental health • quality of life (such as lifestyle and work) • other people (for example, family, friends and employers) • ease of use (for example, tablets rather than injection) • where the treatment has to be used (for example, at home rather than in hospital)

	• any other issues not listed above
Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	• aspects of the condition that the treatment cannot help with or might make worse • side effects (for example, type or number of problems, how often, for how long, how severe. Please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate) • Are other TKI'S available better.
Patient population	
13. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Patients with advanced (stage 3 or 4) disease are likely to require TKI's to extend their life. which introducing Cabozantinib as a first line drug is another option.
Equality	
14. Are there any potential equality issues that should be taken into account when	**None known**

considering this condition and the technology?	
Other issues	
15. Are there any other issues that you would like the committee to consider?	
Topic-specific questions	
16. [To be added by technical team if required, after receiving the company submission. For example, if the company has deviated from the scope (particularly with respect to comparators) – check whether this is appropriate. Ask specific, targeted questions such as “Is comparator X [excluded from company submission] considered to be	

established clinical practice in the NHS for treating [condition Y]?"] **if not delete highlighted rows and renumber below**	

Key messages

17. In up to 5 bullet points, please summarise the key messages of your statement:

- People with advanced kidney cancer have very few treatment options and require a variety of drug choices.
- Cabozantinib can have an acceptable and improved side effect profile compared to other first line drugs, which will improve people's quality of life and hopefully extend a patient's life.
- The future will hopefully be more development in immunotherapy as second line and hopefully there will be better outcomes in survival rates and a better quality of life for patients living with advanced kidney cancer. So keeping patients on a first line treatment for as long as possible gives patients hopefully more time to live.

Different drugs work for different people. Cabozantinib acts in a similar way to the other TKI's but has been shown to target c-MET and VEGFR2 tyrosine kinases. So this will suit certain patients diease better.

- It is vital that people with advanced renal cancer can live for a long time and have a variety of drugs available to keep them living with renal cancer.
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

NICE National Institute for Health and Care Excellence

Clinical expert statement

Cabozantinib for untreated metastatic renal cell carcinoma [ID1208]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**Brighton and Sussex University Hospitals**

3. Job title or position	**Consultant Medical Oncologist**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ✓ a specialist in the treatment of people with this condition? ✓ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	✓ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	When used in the first line management of intermediate and poor prognosis metastatic renal cancer, the aim with Cabozantinib is to improve the likelihood of a patient obtaining a clinical benefit from treatment, improve the duration of benefit from treatment and potentially improve overall survival. Within these, the aim is to deliver the benefits without an increase in overall treatment toxicity compared to existing therapy.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Treatment responses and benefits can occur via either a formally defined complete or partial response or by producing a period of disease stability. These benefits are associated with improvements in patient's symptoms and quality of life.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Intermediate and poor risk metastatic renal cancer has been generally treated with the TKI drugs Sunitinib and Pazopanib since 2009. Prior to their introduction there was no routine beneficial therapy. Whilst many patients obtain significant benefit from these therapies, the overall results are however relatively modest. Some patients, particularly those with poor risk disease, are too unwell to go on to receive second line therapy and a more effective first line therapy would be of potential greater benefit in these patients.
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	The large majority of patients in the UK with intermediate or poor risk metastatic renal cancer are treated with first line TKI therapy. There is a choice of Sunitinib or Pazopanib which are similar drugs in terms of efficacy but have differing side effect profiles.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Management Guidelines for Renal Cancer are available from ESMO, EAU, NCCN and other organisations.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The management of intermediate and poor risk metastatic renal cancer is very similar across the UK. There is a choice of two similar drugs of the current first line therapy (Sunitinib or Pazopanib), aside from this care is very similar between centres and individual clinicians.
• What impact would the technology have on the current pathway of care?	If Cabozantinib were used as first line therapy the current pathways would remain unchanged.
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Cabozantinib is already routinely used in the NHS as a second line therapy for metastatic renal cancer. The drug has a similar mechanism of action and side effect profile to the TKI drugs, Sunitinib and Pazopanib, used currently as the first line therapies.

• How does healthcare resource use differ between the technology and current care?	The use of Cabozantinib rather than Sunitinib or Pazopanib as the first line therapy would not affect the level of healthcare resource use but would increase the duration.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Cabozantinib would be available through specialist oncology clinics.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	No additional investment or training would be needed
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	
• Do you expect the technology to increase length of life more than current care?	The data from the CABOSUN trial indicated a non-statistically significant trend towards increased survival for patients treated with first line Cabozantinib. The CABOSUN trial was relatively small and many patients received additional lines of therapy, this makes a clear conclusion regarding the impact on survival difficult at present.

• Do you expect the technology to increase health-related quality of life more than current care?	The side effect profiles for Cabozantinib and Sunitinib/Pazopanib are very similar and changes in quality of life are likely to be modest.
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	The benefits of Cabozantinib appear to cover the whole patient group. However, a group of patients who have a greater gain are those who are unwell at diagnosis and in the event of progression on Sunitinib/Pazopanib are too ill to receive second line therapy with Cabozantinib or Nivolumab. In these group receiving a drug with a higher rate of clinical benefit and duration of benefit would be advantageous.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors	The practical issues with using Cabozantanib for first line treatment are almost interchangeable with those already in place for the current treatment with Sunitinib/Pazopanib. One minor difference is the need to routinely check and occasionally treat low serum Magnesium levels in patients treated with Cabozantinib.

affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	The practical management including commencing, monitoring and stopping treatment for patients treated with Cabozantinib would be essentially identical to those already established for treatment with Sunitinib/Pazopanib.
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No
17. Do you consider the technology to be innovative in its potential to make a significant and substantial	Cabozantinib is a newer and potentially more powerful TKI therapy. This form of therapy has been the routine management of metastatic renal cancer for the past decade. It would be regarded as a clinically useful improvement to the technology rather than a new innovation.

impact on health-related benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes
• Does the use of the technology address any particular unmet need of the patient population?	The clinical benefits in terms of enhanced response rates and duration of treatment benefit from Cabozantinib appear to apply across the patient population.
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The CABOSUN trial did not include formal quality of life data, however the side effect profiles of Cabozantinib and Sunitinib in the study appear very similar. Cabozantinib is already widely used in the second line management of renal cancer and most clinicians will be familiar with the toxicity profile and management.
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	The CABOSUN trail compared Cabozantinib with Sunitinib in the first line therapy of intermediate and poor risk metastatic renal cancer. However, in current UK practice approximately 50% of this group of patients would currently receive Pazopanib rather than Sunitinib.
• If not, how could the results be extrapolated to the UK setting?	A number of studies have compared the tolerability and efficacy of Sunitinib and Pazopanib and have found no significant difference in efficacy between these two drugs. As a result the CABOSUN data could be extrapolated against either current first line drug.
• What, in your view, are the most important outcomes, and were they measured in the trials?	The CABOSUN study measured response rates, progression free survival and treatment toxicity. Additionally, data on the impact of the choice of first line therapy on overall survival was included.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	N/A
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
20. Are you aware of any relevant evidence that might	No

not be found by a systematic review of the trial evidence?	
21. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]?	There have been a number of studies that give further evidence on the efficacy and tolerability of Sunitinib. These would include the COMPARZ and PISCES trials and real-world database studies such as published by Ruiz-Morales et al.
22. How do data on real-world experience compare with the trial data?	The real world and trial data are similar but the real world has more unwell, poor performance and elderly patients and slightly less impressive data.
Equality	
23a. Are there any potential equality issues that should be taken into account when considering this treatment?	No equality issues are apparent from a change from Sunitinib/Pazopanib to Cabozantinib in this treatment setting.

23b. Consider whether these issues are different from issues with current care and why.	N/A
Topic-specific questions	
24. [To be added by technical team if required, after receiving the company submission. For example, if the company has deviated from the scope (particularly with respect to comparators) – check whether this is appropriate. Ask specific, targeted questions such as "Is comparator X [excluded from company submission] considered to be established clinical practice in	

the NHS for treating [condition Y]?"] **if not delete highlighted rows and renumber below**	
Key messages	
25. In up to 5 bullet points, please summarise the key messages of your statement. • Important gains in response rate and progression free survival • Potential survival gains • Cabozantinib is already widely used in the second line therapy of renal cancer • Similar side effects and treatment plan to the current therapy • A ten-year update/improvement on Renal Cancer TKI therapy	

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

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

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors 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 Chorozoglou, Senior Research Fellow (Health Economics)

Correspondence to Dr Jonathan Shepherd
Southampton Health Technology Assessments Centre (SHTAC)
Wessex Institute
Alpha House
Enterprise Road, University of Southampton Science Park
Southampton SO16 7NS
www.southampton.ac.uk/shtac

Date completed: 10th April 2018

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 17/109/10.

Acknowledgements

We are grateful to the clinical advisors to the ERG:
Professor Robert Hawkins, Cancer Research UK Professor/Honorary Consultant in Medical Oncology, Christie Cancer Hospital Manchester and University of Manchester.
Dr Matthew Wheater, Consultant Medical Oncologist, University Hospital Southampton NHS Foundation Trust.

We would also like to thank: Karen Welch, Information Specialist at SHTAC, for appraising the literature search strategies in the company's submission, running updates of the company's clinical effectiveness searches and searching for ongoing studies; and Dr Emma Loveman Senior Reviewer / Partner, Effective Evidence LLP, for providing a quality assurance review of the draft ERG report.

Declared competing interests

None from the authors. Professor Hawkins declares receipt of speaker fees/honoraria from Ipsen, Novartis, Pfizer and Aveo pharmaceuticals within the previous 12 months. Dr Jenner declares being paid as a speaker for a Pfizer sponsored evening meeting on the topic of managing co-morbidities in metastatic RCC patients (July 2017).

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:

Shepherd, J., Lord, J., Onyimadu, O., Harris, P., Frampton, G., Chorozoglou, M. Cabozantinib for untreated locally advanced or metastatic renal cell carcinoma: A Single Technology Appraisal. Southampton Health Technology Assessments Centre (SHTAC), 2018.

Contributions of authors

Jonathan Shepherd critically appraised the clinical effectiveness review, drafted the report, project managed the assessment and is the project guarantor. Joanne Lord critically appraised the economic evaluation, and drafted the report. Petra Harris critically appraised the clinical effectiveness review and drafted the report. Olu Onyimadu critically appraised the health economic review, critically appraised the economic evaluation, and drafted the report. Geoff Frampton critically appraised the clinical effectiveness review and drafted the report. Maria Chorozoglou critically appraised the economic evaluation, and drafted the report.

Word count: 45,992

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AE	Adverse event
AJCC	American Joint Cancer Committee
CABOSUN	Cabozantinib-sunitinib
CHMP	Committee for Medicinal Products for Human Use
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
DA	Dynamic allocation
DIC	Deviance information criteria
EAU	Updated European Association of Urology
ECOG	Eastern Cooperative Oncology Group
EPAR	European Public Assessment Report
ERG	Evidence Review Group
ESMO	European Society of Medical Oncology
FDA	Food and Drug Administration
FP	Fractional polynomial
HR	Hazard ratio
HRQoL	Health related quality of life
ICER	Incremental cost effectiveness ratio
IMDC	International Metastatic RCC Database Consortium
IRC	Independent radiology committee
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenous
KM	Kaplan-Meier
MET	Hepatocyte growth factor receptor protein
MCMC	Memorial Sloan-Kettering Cancer Center
NCCN	National Comprehensive Cancer Network
NE	Not estimable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
OS	Overall survival
ORR	Objective response rate
PAS	Patient access scheme
PD	Progressed disease
PF	Progression free
PFS	Progression free survival
PH	Proportional hazards
PPES	Palmar-plantar erythrodysaesthesia syndrome
QALY	Quality adjusted life year
QoL	Quality of life
RCC	Renal cell carcinoma
RTKs	Receptor tyrosine kinases
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
TEAE	Treatment emergent adverse events
TNM	Tumour Node Metastasis
TTD	Time to treatment discontinuation
VEGF	Vascular endothelial growth factor (VEGF)

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 recommend 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 (≥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 ≥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

21.0 months (95% CI 16.3, 27.0) in the sunitinib arm, with a median difference of 9.3 months. The data were immature at this data cut-off, and there was a notable degree of censoring around the median estimates (censoring due to no event as of the cut-off date: cabozantinib 52%, sunitinib 42%). Hence, the data should be interpreted with caution.

- The HR, stratified by IMDC risk category and bone metastases, was 0.74 (95% CI 0.47, 1.14) p=0.1700.
- The percentage of patients event-free at 30 months was 50.7% for cabozantinib and 30.3% for sunitinib.
- Updated OS results at the 1st July 2017 data cut off: median OS for cabozantinib was 26.6 months (95% CI 14.6, not estimable) versus 21.2 months (95% CI 16.3, 27) for sunitinib. The HR was 0.80 (95% CI 0.53, 1.21) 2-sided p-value = 0.29. These data are also immature.

ORR

The ORR was 20% (95% CI 12.0%, 30.8%) in the cabozantinib arm, compared to 9% (95% CI 3.7%, 17.6%) in the sunitinib arm. The difference between groups in ORR was 11.3% (95% CI, 0.4 22.2%; p=0.0406). The ORR was classed as a 'confirmed partial response'. There were no confirmed complete responders in either study group.

Subgroups

There was a consistently favourable effect on PFS for cabozantinib compared with sunitinib in pre-defined subgroups (e.g. age, sex, race, baseline ECOG status, bone metastases). Confidence intervals were wide and included 1 for some the smaller subgroups. Subgroup results for OS also showed a favourable effect for cabozantinib compared with sunitinib, however, in most subgroups the confidence intervals included 1. Caution is advised given the observational nature of subgroup data and small sample sizes.

Adverse events

The majority of patients had at least one treatment-related adverse event regardless of treatment arm (95%-97%). Around half experienced a serious adverse event (49%-51%) and just over a third of all patents had a treatment-related serious adverse event (36%). Over half of all patients experienced a Grade 3 or 4 adverse event (60%-63%).

Discontinuations of study drug due to adverse events was also similar between study groups (21%-22%). Patients receiving cabozantinib had longer treatment exposure compared to those receiving sunitinib (median: 6.5 months versus 3.1 months, respectively) and dose

reductions were frequent with both treatments (46% and 35%, respectively), as were dose interruptions (73% and 71%, respectively).

The percentage of patients dying up to 30 days after last dose of study treatment was higher in the sunitinib group compared to the cabozantinib group (11% versus 5.1%, respectively), as was the case for death > 30 days after last dose of study treatment (49% versus 44%, respectively).

The most common adverse events (of any grade) in the cabozantinib treatment group were diarrhoea (72%), fatigue (62%), aspartate aminotransferase increased (60%), hypertension (56%), alanine aminotransferase increased (54%), decreased appetite (45%) and palmar-plantar erythrodysaesthesia syndrome (42%). In the sunitinib group common adverse events were fatigue (67%), platelet count decreased (58%), diarrhoea (49%), anaemia (44%) hypertension (38%), nausea (36%) and neutrophil count decreased (35%).

Indirect treatment comparison

The company conducted indirect treatment comparisons (ITCs) to compare cabozantinib against pazopanib given the lack of head-to-head evidence for these two treatments. The company's ITCs include two RCTs: CABOSUN (cabozantinib versus sunitinib) and COMPARZ (sunitinib versus pazopanib). The comparison with pazopanib is made through the common comparator sunitinib.

Due to the company's observation that proportional hazards do not hold for all survival outcomes in both trials the company used two Bayesian statistical ITC methods that do not assume proportionality in hazards:

- The parametric survival curve method by Ouwens et al provides survival estimates for a family of parametric distributions (Weibull, log-logistic, log-normal, Gompertz, exponential) and can extrapolate outcomes as described by two parameters (shape and scale);
- The fractional polynomial method by Jansen provides survival estimates for first order and second order models from a set of powers (five models for each order, 10 models in total). From these 10 models a best-fitting model was chosen by the company (second order P1=-1 and P2=-1) based on the deviance information criteria.

Both the Ouwens et al and fractional polynomial methods provide OS and PFS effect estimates that are used in the company's economic model.

The CS reports the results of the ITC as fitted survival curves for the outcomes of OS and PFS for all three treatments (cabozantinib, sunitinib, pazopanib), based on fixed effect and on random effects, for each of the five parametric distributions generated by the Ouwens et al method. For each of the analyses cabozantinib had a higher survival estimate than sunitinib or pazopanib. The sunitinib and pazopanib curves were similar to each other in shape and position, indicating similar effectiveness between these two treatments.

The CS presents fitted fractional polynomial survival curves for the outcomes of OS and PFS for all three treatments, based on fixed effects for first and second order models. On request the company also supplied HR plots with credible intervals for each fractional polynomial model to allow visual inspection of the time-varying HR curves. Results for PFS from the best-fitting fractional polynomial model (which informs the economic model base case) show:

- The HR for pazopanib peaks at month four [redacted] and declines slightly during the rest of the follow-up period. The HR for sunitinib peaks at month six [redacted] and declines slightly during the remainder of the follow-up period.
- The credible intervals increase over the follow-up period, with the upper bound increasing to include 1 after month 19 for pazopanib, and after month 11 for sunitinib.
- The time-varying PFS HRs for cabozantinib versus sunitinib generated by this fractional polynomial model compare broadly with the constant HR reported in the CABOSUN trial (0.48 (95% CI 0.31, 0.74)), though with greater uncertainty (wide credible intervals).

Results for OS from the best-fitting PFS fractional polynomial model (which informs the economic model base case) show:

- The HR for pazopanib starts to peak at month nine, and declines slightly after month 19 [redacted] The HR for sunitinib begins to plateau at month 13 and peaks at month 30 where it remains for the rest of the follow-up period [redacted].
- The credible intervals widen during the course of the follow-up period, and include 1 at all time points.
- The time-varying OS HRs for cabozantinib versus sunitinib generated by this fractional polynomial model compare broadly with the constant OS HR reported in the CABOSUN trial (0.80 (95% CI 0.53, 1.21), though with greater uncertainty (wide credible intervals).

Across the other fractional polynomial models (first and second order), the time-varying HR curves for cabozantinib versus sunitinib and cabozantinib versus pazopanib have a similar

shape to each other. Cabozantinib is of superior effectiveness when compared with both sunitinib and with pazopanib, with little difference between the results of each pairwise comparison.

The ERG considers that the statistical methods used to conduct the ITC are appropriate, but there is uncertainty in the results due to differences between the trials in patient prognostic characteristics (more detail on the critical appraisal of the ITC is available below under 'Commentary on the robustness of submitted evidence').

Summary of submitted cost effectiveness evidence

The CS includes:

- A review of published cost-effectiveness studies relating to cabozantinib, sunitinib and pazopanib in previously untreated locally advanced or metastatic RCC.
- An economic evaluation undertaken for the NICE STA process, comparing cabozantinib with pazopanib and sunitinib in treatment-naïve patients with locally advanced or metastatic RCC.

The company conducted a systematic search of the literature to identify economic evaluations with cabozantinib, sunitinib or pazopanib in untreated advanced RCC. The search identified 23 published cost-effectiveness studies, of which seven were conducted from an English, Welsh or British perspective. The company concluded that as none of the studies included cabozantinib, they are not directly relevant to this appraisal.

The company developed a model to evaluate the cost-effectiveness of cabozantinib as first-line treatment for advanced RCC. The model is a health state transition model, containing three mutually-exclusive health states: progression free (PF); progressed disease (PD) and death. Patients start in the PF state, and at disease progression, transition to the PD state, which is considered irreversible. Patients in PF and PD states die from cancer or other causes.

The distribution of the cohort between the health states and treatment states at each time point is estimated using a partitioned survival approach, based on PFS, Time to Treatment Discontinuation (TTD) and OS curves:

- **Death**: The proportion of patients alive at each time point is taken from the OS curve. Hence, the proportions of the cohort who die in each cycle are calculated.

- **PF**: The proportion of patients who are progression free is the minimum of the PFS curve and the OS curve at each time point.
- **PD**: The proportion of patients in the PD state is calculated as the residual (if any) of the cohort who are not dead and not progression free.

Patients enter the PF state on first-line treatment but may stop at any time due to adverse effects or when their disease progresses. After a fixed waiting period of 8 weeks, most patients then progress to subsequent treatment with one of 10 drugs included in the company's base case. The duration of second-line treatment is defined for each drug, after which patients are assumed to receive supportive care until death. The proportion of patients on first-line treatment is determined by the minimum of the TTD and PFS curves. Subsequent treatment status is calculated based on a waiting time and defined treatment duration for each individual second-line drug.

Other key features and assumptions of the model are listed below:

- **Cycle length**: 1 week, with half cycle correction.
- **Time horizon**: 20 years in base case (with 10 years in scenario analysis).
- **Duration of treatment effects**: based on extrapolation of PFS and OS curves fitted to trial data, assuming no waning of benefits over the time horizon.
- **Adverse events**: For each first-line treatment, grade 3 or 4 Treatment Emergent Adverse Events (TEAEs) with an incidence of 5% or more are included in the model. There is no explicit modelling of adverse events related to subsequent treatments.
- **Utility and QALY calculations**: Utility weights for the PF and PD health states are based on published estimates, assumed independent of treatment. Additional disutilities are applied to reflect included TEAEs for first-line treatments – applied as a one-off QALY loss in the first cycle. QALYs are also adjusted for the gender mix and age of the cohort.
- **Health resource use and costs**: The model estimates costs for of first-line and subsequent treatment; monitoring and disease management in PF and PD states; treatment of TEAEs for first-line treatments; and end of life care, applied in the last cycle before death.
- **Discounting**: 3.5% per year for costs and QALYs.
- **Uncertainty**: the model allows for exploration of uncertainty over input parameters using deterministic sensitivity analysis; scenario analyses varying selected model assumptions; and probabilistic sensitivity analysis (PSA) to estimate the joint effects of parameter uncertainty on the estimated costs and QALYs.

To apply the partitioned survival model, OS, PFS and TTD curves are required for cabozantinib and comparators, extrapolated over the 20-year time horizon. The company present two sets of base case results:

1. **Direct comparison** (cabozantinib versus sunitinib)
   This analysis is based on patient-level data from the CABOSUN trial, with OS, PFS and TTD curves separately fitted for cabozantinib and sunitinib arms using six families of survival functions: exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma.
2. **Indirect comparison** (cabozantinib versus sunitinib and pazopanib)
   The two ITC meta-analyses methods described above were used to estimate PFS and OS curves:
   - ITC parametric curves, fixed and random effect models for five survival functions: exponential, Weibull, Gompertz, log-normal, log-logistic. The generalised gamma distribution was not implemented due to the lack of the incomplete gamma function in WinBUGS. The company reports that treatment was tested as a covariate, but the model only includes curves that were fitted separately for cabozantinib and sunitinib.
   - ITC fractional polynomial curves, with five first-order and five second order functions.

   As TTD data are not available from COMPARZ, the company uses curves fitted to CABOSUN for cabozantinib and sunitinib, and assumes that the latter also apply to pazopanib. This assumption is justified by the similar median and mean duration of treatment between the treatment arms in COMPARZ.

Survival curves for PFS, OS and TTD used in the company's base case and scenario analyses are summarised in the table below.

Table 1 Survival curves used in company analyses

Curve	Method	Treatment	CS Base case	CS scenarios
PFS	**Direct** CABOSUN	Cabozantinib Sunitinib	Log-normal	Exponential Weibull Gompertz
	ITC CABOSUN & COMPARZ	Cabozantinib Sunitinib Pazopanib	FP P1=P2=-1	RE exponential RE Weibull RE Gompertz
OS	**Direct** CABOSUN	Cabozantinib Sunitinib	Exponential	Exponential Weibull Gompertz
	ITC CABOSUN & COMPARZ	Cabozantinib Sunitinib Pazopanib	FP P1=P2=-1	RE exponential RE Weibull RE Gompertz FP P1=-0.5, P2=0 FP P1=-1, P2=0
TTD	**Direct** CABOSUN	Cabozantinib Sunitinib Pazopanib	Log-normal	Exponential Weibull Gompertz Generalised gamma

PFS, progression free survival; OS, overall survival; TTD, time to treatment discontinuation; ITC, indirect treatment comparison; RE, random effects; FP, fractional polynomial

Base case utility estimates for the PF and PD health states were taken from a previous NICE appraisal (Tivozanib TA512). The company also tested scenarios using health state utility estimates from the NICE appraisals for sunitinib and pazopanib, and the Swinburn et al 2010 study. These utility estimates are shown in the table below.

Table 2 Utility values (adapted from CS Tables 46 & 47)

Health state	Utility value: mean (SE)	95% CI	CS reference
Progression free			
Base case	0.726 (0.011)	0.705 to 0.748	Tivozanib TA512
Scenario	0.70 (0.01)	0.680; 0.720	Pazopanib TA215
Scenario	0.78 (0.078)*	0.627; 0.933*	Sunitinib TA169
Scenario	0.795 (0.0176)	0.761; 0.830	Swinburn 2010
Progressed disease			
Base case	0.649 (0.019)	0.612 to 0.686	Tivozanib TA512
Scenario	0.59 (0.059)*	0.474; 0.706*	Pazopanib TA215
Scenario	0.705 (0.071)*	0.567; 0.843*	Sunitinib TA169
Scenario	0.355 (0.0288)	0.299; 0.412	Swinburn 2010
TEAE grade ¾			
Base case	-0.2044 (0.0682)	-0.0707 to -0.3381	COMPARZ
Scenario	-0.0550 (0.0068)	-0.0418; -0.0685	METEOR trial

Abbreviations: TEAE, treatment emergent adverse effects; CI, confidence interval; SE, standard error; *SE or 95% CI not available in literature; 10% of the mean assumed.

Utility estimates for adverse events are sourced from the COMPARZ (base case analysis) and METEOR (scenario analysis) trials (METEOR was a phase III trial which compared cabozantinib with everolimus in patients with RCC that had progressed after VEGFR-targeted therapy[51]). The company assumes that the utility effects of adverse events are not disease-specific and that all types of grade 3 or 4 events elicit the same utility loss for a fixed period of 4 weeks and a fixed number of episodes per patient per TEAE. These assumptions yield a mean QALY loss of 0.0225 per TEAE in the base case (0.006 in the METEOR trial-based scenario). The company models the incidence of grade 3/4 TEAEs based on reported rates from the CABOSUN study for cabozantinib and sunitinib and from COMPARZ for pazopanib. Only events with a reported incidence of 5% or greater in at least one arm were included.

The company conducted a systematic literature review to identify published resource use and cost data relevant to the cost-effectiveness analysis. The costs included in the economic model are acquisition and administration of first-line and subsequent treatments, with adjustment for dose intensity and wastage when appropriate; monitoring and disease management in PF and PD states; treatment of TEAEs for first-line treatments; and end of life care. The CS reports PAS prices for first-line treatment but not second-line treatment. The company consulted UK-based clinical experts for the estimation costs accruing from of health state management resources, adverse event resource use and end-of-life care.

The results of the economic model are presented as incremental cost effectiveness ratios (ICERs).

- For the company base case, using the direct comparison from the CABOSUN trial, an ICER of £37,793 per QALY gained is reported for cabozantinib versus sunitinib.
- Based on their preferred ITC model, sunitinib is dominated by pazopanib and the ICER for cabozantinib compared with pazopanib is £48,451. The pairwise ICER for cabozantinib compared with sunitinib in this model is £31,538.
- Probabilistic results were similar.

The company conducted one-way deterministic sensitivity analyses and concluded that the key drivers to the cost-effectiveness results include drug costs and discount rates for QALYs and costs. Other parameters identified in the company's one-way sensitivity analysis include relative dose intensity and utilities associated with the progression-free state. The company's scenario analyses found cost-effectiveness results to be most sensitive to the choice of OS curve used in the model.

Commentary on the robustness of submitted evidence

Strengths

- The literature searches conducted by the company were considered by the ERG to be appropriate and sufficiently comprehensive to have identified all the relevant clinical effectiveness evidence. The company's systematic review methods were considered appropriate.
- The CABOSUN trial provided a direct comparison with sunitinib for PFS, OS, tumour response and adverse events. The ERG considers it to be well conducted overall, though there is a lack of detail on randomisation and concealment of allocation procedures to inform assessment of risk of selection bias. The open-label nature of the trial means the potential risk of performance and detection bias. However, the retrospective blinded IRC assessment of tumour response and progression conducted for the regulatory submission reduces the risk of detection bias for PFS and tumour response outcomes. This trial has some further limitations as described below.
- The Ouwens and fractional polynomial ITC methods appear to have been implemented adequately in accordance with the original methodological publications and the ERG considers that both are suitable for use for the indirect comparison of treatments in this appraisal. However, the results of both methods may be biased by the differences in RCC risk factors and other variables between the CABOSUN and COMPARZ trials (see below).
- The company's systematic review of cost effectiveness was of good methodological quality. The ERG agrees with the company's conclusion that none of the studies identified from the literature review included cabozantinib, and as such, they are not directly relevant to this appraisal.
- The two studies (CABOSUN and COMPARZ) used to estimate outcomes of PFS and OS provide the best available data sources, although the ERG does have concerns about the differences in patient population. The company conducted a range of ITC curve fitting methods (parametric and fractional polynomial methods) and used the resulting curves to make the indirect comparison from cabozantinib to pazopanib and to extrapolate beyond the trial follow up.
- The structure of the company's model reflects the nature of progression and clinical pathway for people with previously untreated locally advanced or metastatic renal cell

cancer. The company used methods for the economic evaluation that are consistent with NICE methodological guidelines and with other drug appraisals for this population.

- The ERG agrees that the health state utility values applied in the company's model meet the NICE reference case and are suitable for inclusion in the model. Costing methods and sources are also generally of good standard with reasonable assumptions. Scenario analysis reflective of the current NHS practice are explored.

Weaknesses and areas of uncertainty

- The CABOSUN trial has some limitations.
    - It is a phase II trial, and was never designed to be a registration trial. It has a relatively small sample size (n=157 patients).
    - The trial was conducted entirely in the US and therefore it may not necessarily be applicable to the UK (though clinical experts to the ERG regarded the baseline characteristics as generally representative of patients in their practice).
    - OS was a secondary outcome and the data are immature. HRQoL was not an outcome measure.
    - The updated PFS assessment conducted for the regulatory submission (and used in the CS) used different censoring rules and a blinded IRC, which meant that the number of events (progressions or deaths) recorded (n=92) was less than the number required in the original PFS statistical power calculation (n=123). This means the updated PFS assessment would be statistically under-powered.
    - There was an imbalance between trial arms in the number of patients with missing data. One patient in the cabozantinib arm and six in the sunitinib arm withdrew prior to receiving study treatment, but the reasons for these withdrawals were not known. There was also a higher incidence of missing or unevaluable data in the sunitinib arm, with six patients in the cabozantinib arm and 18 in the sunitinib arm not evaluable because they had no adequate post-baseline imaging assessments. The CS states that based on their baseline characteristics (data unavailable to the ERG to verify), the sunitinib patients without post-baseline imaging would not be expected to have a better prognosis than sunitinib patients who had a response recorded, and therefore it is unlikely that the radiographic endpoints were biased against sunitinib by these missing data.
- There are some important differences between the two trials in the ITC:

    - The CABOSUN trial included only patients at intermediate or poor RCC risk, whilst the COMPARZ trial included patients at favourable, intermediate and poor risk classifications.
    - A greater proportion of COMPARZ patients were classified as having the highest cancer performance status, likely due to inclusion of some patients with favourable RCC risk status in the trial.
    - Around a third of patients in CABOSUN had bone metastases (a key prognostic factor in RCC) at baseline compared to 18% of patients in COMPARZ.
    - The impact of these differences on the results of the ITC are not discussed in the CS. The ERG considers that the impact of the differences on the ITC results to be uncertain.
- The ERG believes that the company's cost effectiveness results include some errors in model inputs and calculations, which could bias conclusions on cost-effectiveness. The ERG corrected errors in the company's QALY calculations and small errors in costs.
- It is appropriate to estimate costs and health effects over the patients' whole lifetimes, so we do not disagree per se with the company's use of a 20-year time horizon. However, other RCC NICE appraisals have adopted a more conservative time horizon of 10 years. In the company's base case model, a relatively small proportion of the modelled cohort survive to 10 or 20 years. However, we question the extrapolation of OS and PFS curves from limited trial follow-up over 20 years. This entails strong assumptions about persistence of treatment effects, which may not be realistic. We investigate the impact of the time horizon and different assumptions about waning of treatment effects in ERG analysis.
- Although the company's preferred survival models have reasonable face validity with good measures of fit, they appear to overestimate PFS and OS. We note that other fitted models do not necessarily address this uncertainty. Based on measures of fit and plausibility of extrapolation, the ERG agrees with the company's selection of best direct comparison and ITC parametric and fractional polynomial curves. However, selection of curves for scenario analyses fit less well. We explored alternative assumptions in the ERG scenario analysis. For the direct comparison, we note that the company's choice of the exponential distribution for both cabozantinib and sunitinib conflicts with the conclusion that OS hazards are not proportional. However, we suggest that the exact shape of the CABOSUN Kaplan-Meier (KM) OS 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.

- 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 does not explicitly state which OS dataset was used to inform in the model, but the January 2017 KM plot is reproduced in the CS economic chapter and KM data provided by the company in response to a clarification question also relates to this earlier cut-off). 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, we assume equivalent OS for pazopanib and sunitinib, based on the results of COMPARZ.
- **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

case, we estimate that sunitinib is dominated by pazopanib due to its higher cost and similar effectiveness.

Table 3 Cost-effectiveness: ERG preferred assumptions analysis results

Drug	Costs (£)	QALYs	Life-years	PF life years	ICER (£ per QALY gained)	
					Incremental analysis	Pairwise, cabozantinib vs. comparator
Pazopanib	[redacted]	[redacted]	[redacted]	[redacted]	-	65,743
Sunitinib	[redacted]	[redacted]	[redacted]	[redacted]	-	41,465
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	65,743	-

However, this result was sensitive to some cost and resource use assumptions. By assumption, our preferred analysis gave the same life expectancy with sunitinib as with pazopanib, yielding very similar QALY estimates. Cabozantinib has a modest survival advantage and a larger effect on progression free survival and hence QALYs. We believe that these results appropriately reflect evidence from CABOSUN and COMPARZ. The results were generally robust, with the ICERs remaining above £30,000 per QALY gained for all of the scenarios that we tested.

The above analyses include existing PAS discounts for cabozantinib, sunitinib and pazopanib for first-line treatments. However, they exclude these arrangements and other existing PAS discounts for subsequent treatment after failure of first line treatment. We present results for the ERG-corrected company base case and scenarios and for ERG additional analysis in a confidential addendum to this report.

The ERG is 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. This is because the submitted CS model and results from the ERG's preferred assumptions give mean OS estimates exceeding 24 months for sunitinib and pazopanib.

1 Introduction to ERG Report

This report is a critique of the company's submission (CS) to NICE from Ipsen Ltd UK on the clinical effectiveness and cost effectiveness of cabozantinib for untreated locally advanced or metastatic renal cell carcinoma (RCC). It identifies the strengths and weakness of the CS. Clinical experts were consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the CS was requested from the company by NICE and the ERG on 22nd February 2018. A response from the company via NICE was received by the ERG on 9th March 2018 and this can be seen in the NICE committee papers for this appraisal.

2 BACKGROUND

2.1 Critique of company's description of underlying health problem

The ERG considers that generally the CS provides a clear and accurate overview of the nature and clinical consequences of renal cell carcinoma (RCC).

2.1.1 Renal cell carcinoma (RCC)

RCC is a cancer that usually originates in the lining of the proximal renal tubules of the kidney - the smallest tubes inside the nephrons that help filter the blood and make urine. As stated in the CS, this is the most common type of kidney cancer, accounting to around 80% of all kidney cancer cases.[1] The three main types of RCC are clear cell (75%), papillary (10%) and chromophobe (5%).[2]

As stated in the CS, early small RCC tumours are usually asymptomatic and are often discovered incidentally during other investigations.[3] [4] In consequence, many patients present with advanced disease (around 38%[5]) and 25-31% of patients present with metastases at diagnosis.[2] The NICE final scope covers both locally advanced RCC (which cannot be removed by surgery) and metastatic RCC.

2.1.2 Clinical presentation

The most common symptom of RCC is blood in the urine (in around 50% of cases).[2] As described in the CS, other non-specific symptoms include weight loss, fever, sweating, fatigue and anaemia amongst others. At the metastatic stage, the tumour has spread beyond the regional lymph nodes to other parts of the body. Less frequent symptoms related to the metastatic spread of the disease include bone pain, skeletal-related events and

hypercalcaemia, as well as venous thromboembolism and lung symptoms such as airway obstruction.[3 6]

2.1.2.1 Staging and prognosis

A staging system is used to show how far the cancer may have spread (and whether it has spread into nearby lymph nodes or distant organs) on a scale of I to IV. Lower stage cancers are less likely to spread than higher stages cancers.[2] The NICE scope denotes stage IV (metastatic) cancer.

One of the most common staging systems (the extent of the cancer in the body) used is the American Joint Cancer Committee (AJCC) Tumour Node Metastasis (TNM) system. This classifies the size of the tumour. In addition to this, the CS presents the International Metastatic RCC Database Consortium (IMDC) risk stratification model (also known as 'the Heng model'). This is the method specified in the NICE final scope and cited in the Summary of Product Characteristics (SmPC). The IMDC is an update of a previous classification system known as the Memorial Sloan Kettering Cancer Center (MSKCC) model.[7 8] The MSKCC model is similar to the IMDC criteria, with the latter was a minor revision of the former.

According to expert clinical advice received by the ERG, the IMDC classification is not formally used in clinical practice in the UK. Level of risk has traditionally been judged on general clinical assessment and blood tests. As newer drug therapies are introduced targeted at specific risk groups use of the IMDC will likely increase. Clinical expert advice to the ERG indicates that the use of the IMDC classification would not require any significant changes to clinical practice. In this method, patients are assessed for the presence of six risk factors (routinely collected in practice):

- Karnofsky performance status (PS) <80%
- Haemoglobin <lower limit of normal
- Time from diagnosis to treatment of <1 year
- Corrected calcium above the upper limit of normal (ULN)
- Platelets greater than the upper limit of normal
- Neutrophils greater than the upper limit of normal

Based on the six risk factors, patients are categorised into three risk groups, which predict survival and influence the management of the patient's RCC:[7]

- Favourable – 0 factors
- Intermediate - 1 or 2 factors
- Poor - >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).[7] 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.[2]

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.[9] 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.[10]

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.[10] [11] As might be expected, evidence shows that the effects of disease progression in these patients is linked to a deterioration in HRQoL.[12] [13] [14] [15]

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.[16] RCC is a sub-type of kidney cancer, accounting for around 80% of all kidney cancer cases, as stated above.

Kidney cancer is the UK's seventh most common cancer, accounting for 3% of incident cases. Partly due to increased detection of early-stage tumours, rates in the UK are estimated to increase annually by 1.2%.[2] There were a reported 3319 deaths from kidney cancer in 2015 in England, with no updated figures as yet available for 2016.[16]

Kidney cancer is more predominant in males, described as a 17:10 male:female ratio in the CS. The ERG note that this was reported as 63% vs 37% (male:female) in the CS-cited source.[2] In the UK, it is the sixth most common cancer in men and the 10th in women and incidence increases with age.[2] For both men and women, the highest rates of kidney cancer are in the 85 to 89 age group.[2]

The risk factors reported in the CS are for kidney cancer only and not specific to RCC: 42% major lifestyle and other risk factors, 62% for hypertension, 24% for smoking and 24% for excess bodyweight. However, cigarette smoking, obesity and hypertension are well-established risk factors for RCC.[17] The risk factor of related hereditary syndromes is not reported, most likely because its occurrence is relatively low (approximately 3% to 5%).[18]

2.2 Critique of company's overview of current service provision

The CS provides a generally clear and accurate overview of how locally advanced and metastatic RCC is managed in clinical practice.

Advanced RCC is incurable and largely resistant to chemotherapy, radiotherapy and hormonal therapy. Due to the lack of improved survival with either chemotherapy or hormonal therapy alone, the mainstay of treatment for locally advanced or metastatic RCC starting in the late 1980s were cytokines, of which interferon alfa and interleukin-2 have been the most evaluated.[19] Targeted drug therapies are now the mainstay of treatment, although some patients receive surgery to reduce the size of the tumour or to remove metastases and this may be in addition to drug treatment.[1] The CS states that treatment goals are to extend life, delay disease progression, relieve symptoms and maintain function, citing a previous NICE appraisal (TA178) as reference.[20] Figure 1 illustrates the NICE pathway of care for renal cancer.[21]

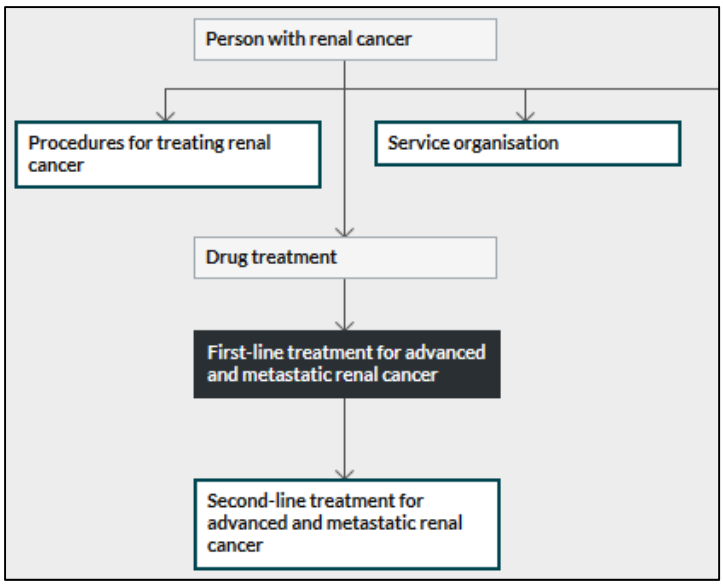


Figure 1 NICE pathway of care in renal cancer

Currently recommended first-line treatments for previously untreated advanced RCC by NICE are:

- Sunitinib in patients who are suitable for treatment and have an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1 (TA169).[12]
- Pazopanib, in patients who have not received prior cytokine therapy and have an ECOG performance status of 0 or 1 (TA215).[13]
- Tivozanib for treating advanced renal cell carcinoma in previously untreated adults (TA512).[22] (NB. This drug was not included in the scope of the current NICE appraisal as its appraisal had not completed at that time).

Bevacizumab, sorafenib and temsirolimus are not recommended by NICE for people with advanced and/or metastatic RCC (TA178).[20]

As there are no UK-specific clinical guidelines for the treatment of RCC, the CS states that in addition to the medicines recommended by NICE, current clinical practice in England and Wales reflects the following guidelines:

- European Society of Medical Oncology (ESMO) Renal Cell Carcinoma: Clinical Practice Guidelines for diagnosis, treatment and follow-up.[1]
- Updated European Association of Urology (EAU) Guidelines: Recommendations for the Treatment of First-line Metastatic Clear Cell Renal Cancer.[23]
- National Comprehensive Cancer Network (NCCN) clinical practice guidelines in oncology, kidney cancer.[6]

2.2.1 Proposed place of cabozantinib in the clinical pathway

The CS states that it is anticipated that cabozantinib (in this indication) will be used in accordance with its marketing authorisation ("treatment of advanced renal cell carcinoma (RCC) in treatment-naïve adults with intermediate or poor risk per IMDC criteria"). The CS proposes that cabozantinib would be an additional treatment option alongside sunitinib and pazopanib in intermediate or poor risk patient groups.

Both the EAU and NCCN guidelines have been updated to include cabozantinib as a treatment option in previously untreated IMDC intermediate and poor risk RCC, while the position of the ESMO guidance on cabozantinib as a treatment option in previously untreated RCC is still unclear.

2.2.2 Potential impact on current service provision

As cabozantinib is another orally administered treatment, the CS states that there is no requirement for a change in current management arrangements or infrastructure. The CS states that testing required to assign patients to IMDC risk groups is carried out as part of routine clinical practice. As stated above, the IMDC criteria are not formally used in clinical practice in the UK, but clinical advice to the ERG is the information required to complete the criteria are routinely collected.

Treatment dose modifications can be managed remotely, without the patient having to attend a consultation in person. Cabozantinib is already recommended by NICE within its marketing authorisation for use in previously treated advanced RCC (TA463[14]) and the ERG therefore agrees that there should be no additional impact on current service provision when used in a previously untreated patient group.

2.3 Critique of company's definition of decision problem

The CS provides a summary table (CS Table 1) including the final decision problem issued by NICE, the company's decision problem and a rationale for any differences between the two.

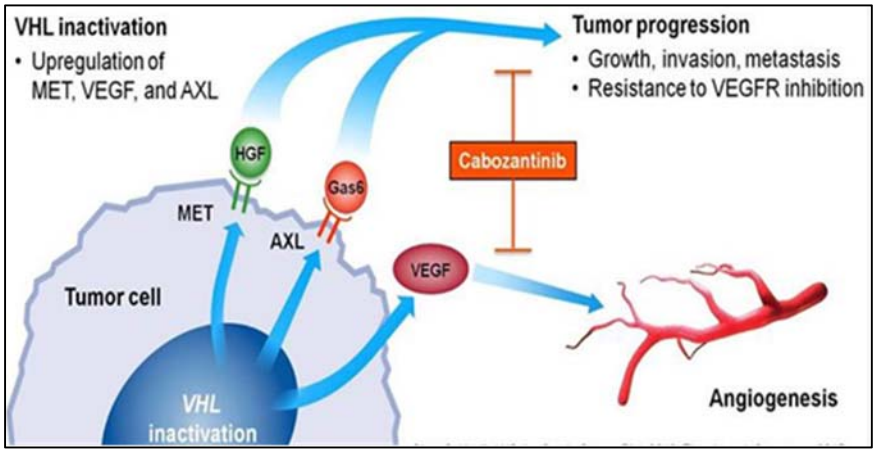
2.3.1 Population

The population specified in the company's decision problem is people with untreated, intermediate or poor risk (as per IMDC criteria), locally advanced or metastatic RCC. The CABOSUN trial (which is the main cabozantinib clinical effectiveness study in the CS) focused on IMDC intermediate- and poor-risk groups, the rationale being that these groups capture 70% to 80% of all patients with advanced disease and because such patients are most in need of systemic therapy and disease control.[24] The patient population matches that

specified in the final scope issued by NICE and that specified in the SmPC indication for cabozantinib (application for marketing authorisation for cabozantinib for "the treatment of advanced RCC in treatment-naïve adults with intermediate or poor risk per IMDC criteria" was submitted to the European Medicines Agency (EMA) on 28 August 2017. Cabozantinib is already licensed for the treatment of advanced RCC in adults following prior VEGF-target therapy.

2.3.2 Intervention

In accordance with the final NICE scope, the intervention described in the company's decision problem is cabozantinib (brand name CABOMETYX®). Cabozantinib is a small molecule that inhibits multiple receptor tyrosine kinases (RTKs) implicated in tumour growth and angiogenesis, pathologic bone remodelling, drug resistance, and metastatic progression of cancer. Figure 2 shows that cabozantinib has a multi-targeted mechanism of action in the treatment of RCC, targeting and inhibiting the MET (hepatocyte growth factor receptor protein), VEGF (vascular endothelial growth factor) and AXL receptors.



Source: CS Section B.1.2 Figure 1

Figure 2 Cabozantinib's mechanism of action

The company supplied the SmPC with their submission to NICE. The revised European Public Assessment Report (EPAR) is not yet available. On March 22nd 2018 the Committee for Medicinal Products for Human Use (CHMP) expressed a positive opinion on the use of cabozantinib for first line treatment of adults with intermediate or poor risk advanced RCC. As outlined in the CS, the SmPC states the recommended dose of cabozantinib is a once-a-day tablet of 60 mg, but is also available as 20 and 40 mg. Treatment should continue until the patient is no longer clinically benefitting from therapy (assessed as tumour progression) or until unacceptable toxicity occurs. Suspected adverse drug reactions may require temporary treatment interruption and/or dose reduction. When dose reduction is necessary, it is recommended to reduce cabozantinib to 40 mg daily and then to 20 mg daily. Dose

interruptions are recommended for management of Common Terminology Criteria for Adverse Events (CTCAE) grade ≥3 toxicities or intolerable grade 2 toxicities. Dose reductions are recommended for events that, if persistent, could become serious or intolerable. Cabozantinib, the intervention described in the decision problem, is appropriate for the NHS and reflects its licensed indication.

2.3.3 Comparators

The two comparators of interest listed in the company's decision problem are those specified in the NICE final scope:

- sunitinib
- pazopanib

These comparators are appropriate for the NHS as they have been recommended for first line use by NICE. As previously stated, both sunitinib[12] and pazopanib[13] are licensed as first-line treatment of advanced and/or metastatic RCC.

2.3.4 Outcomes

The company has listed all but one of the outcomes specified in the NICE final scope in their decision problem:

- overall survival (OS)
- progression-free survival (PFS)
- response rates
- adverse effects (AE) of treatment

The NICE finial scope specified health-related quality of life (HRQoL) as an outcome, but no such data were collected in the single phase II trial (CABOSUN) presented in evidence of the clinical effectiveness of cabozantinib. Hence, HRQoL was not presented as a clinical effectiveness outcome measure in the CS (though HRQoL utility data from other sources are used in the economic model).

2.3.5 Economic analysis

The partitioned survival model used in the CS is considered as one of the standard methods for population-based cancer survival analysis and the method is in line with previous health economic analyses.[13 14] (see section 4 of this report for description and critique of the company's economic evaluation).

2.3.6 Other relevant factors

There were no subgroups of relevance noted in the NICE final scope or the company's scope. Although the scope does not require subgroups to be assessed, the CABOSUN trial included subgroup analyses based on a number of factors, including RCC risk and bone metastases which are of prognostic significance.

The company states that they do not anticipate that the use of cabozantinib will be associated with any equality issues.

3 CLINICAL EFFECTIVENESS

3.1 Critique of company's approach to systematic review

3.1.1 Description of company's search strategy

The CS reports four separate systematic literature searches:

- Clinical effectiveness evidence. The search strategy used in the submission to NICE by the manufacturer of pazopanib in NICE TA215.[13] covering the period 1980 to 2009 was adapted by the company and updated to 28th June 2017 (search strategy reported in CS Appendix D).
- Cost effectiveness evidence. Search period: 1946 to 19th September 2017 (search strategy reported in CS Appendix G).
- Health Related Quality of Life (HRQoL). Original search period: 2006 to July 2016; Update search period: 2016 to 28th July 2017 (search strategy reported in CS Appendix H).
- Cost and healthcare resource identification measurement and valuation. Original search period: 2006-2016; Update search: 2016 to 19th September 2017 (search strategy reported in CS Appendix I).

The clinical effectiveness search strategy was designed from a global perspective. It included search terms for a range of treatments including those within the scope of the appraisal (cabozantinib, pazopanib and sunitinib) and others not in the scope (interferon alfa, interleukin-2, sorafenib, bevacizumab and interferon alfa, temsirolimus, tivozanib). The search terms contain appropriate subject headings together with a good range of truncated free text. An appropriate range of databases was searched: Medline (including In-Process and other non-indexed citations); Embase and the Cochrane Library. A combined search filter was used to identify RCTs, controlled and other trials, meta-analyses and systematic

reviews. The clinical effectiveness search strategy is extensive but visually overcomplicated with various date restrictions applied to different sets of drugs. The ERG notes, however, that the sets are correctly combined and the number of hits (records retrieved) per line is documented for transparency. The search write up offers guidance to the strategy, is thorough and transparent.

Supplementary searching was undertaken in the CS to identify ongoing trials on the National Institute of Health's (NIH) clinical trial registry (www.clinicaltrials.gov) and reference lists from other HTA submission documents were searched. Conference abstracts were not specified as being searched separately. The company state that there are no relevant ongoing studies (CS Section B.2.11), and in response to a clarification request (question A13) stated that they are not aware of any planned or ongoing trials of cabozantinib (as a single therapy agent) for the indication in this appraisal.

The ERG re-ran the company's clinical effectiveness searches on Medline, Embase and the Cochrane Library for the years 2017 to present, to identify any recently published relevant studies. The ERG additionally ran searches of two databases on the Web of Science Platform: Conference Proceedings Citation Index-Science (CPCI-S) and the Emerging Sources Citation Index (ESCI). The following conferences for the years 2016-2018 were additionally searched on the internet: American Society of Clinical Oncology (ASCO), European Society for Medical Oncology (ESMO), European Cancer Organisation (ECCO), European Multidisciplinary Meeting on Urological Cancers (EMUC), International Kidney Cancer Symposium (IKCS). Given that the CS only searched one on-going trials database the ERG checked for any missing ongoing trials on NIHR UKCTG (UK Clinical Trials Gateway), the WHO ICTRP (International Clinical Trials Platform) and re-checked the clinicaltrials.gov database.

Results from these searches were screened by an ERG reviewer. Two relevant conference abstracts not included in the CS were identified, both of which report results from the CABOSUN trial, the sole RCT of cabozantinib included in the company's systematic review of clinical effectiveness (see section 3.1.3 of this report). One of the abstracts[25] was linked to a poster which was included in the CS,[24] albeit the abstract contained less information than the poster. The second abstract[26] identified by the ERG was presented at the European Society for Medical Oncology (EMSO) conference in February 2018 and hence not available at the time the CS was produced. This abstract was linked to a slide presentation and a poster, both of which the ERG were unable to access. The abstract included a small amount

of additional information on tumour response results in a patient subgroup not presented in the CS. We report these data in section 3.3.6 of this report.

The cost effectiveness, HRQoL and health care resource-use searches are much easier to follow as the sets are grouped together more logically, without the varying date ranges. The terms and search filters are all appropriate.

In summary all searches are well documented and are fit for purpose and it is unlikely that any potentially relevant studies comparing cabozantinib with sunitinib and pazopanib were not included.

3.1.2 Statement of the inclusion/exclusion criteria used in the study selection

The company provides a description of the inclusion criteria for the systematic literature review (SLR) (CS B.2.1, Table 4), which was also used to identify studies for potential inclusion in an indirect treatment comparison (ITC) (see Section 3.1.7). These criteria were broader than the NICE final scope but the treatments were subsequently limited to cabozantinib, sunitinib and pazopanib for this appraisal. Details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised are contained in CS appendix D. The interventions and comparators reflect the nature of the decision problem, the anticipated licensed indication and current NHS practice. The CS provides a flow diagram illustrating the number of records identified through each of search sources: the electronic database search, the pazopanib company submission to NICE for TA215[13] and through study registry searches (clinicaltrials.gov) (CS Figure 3).

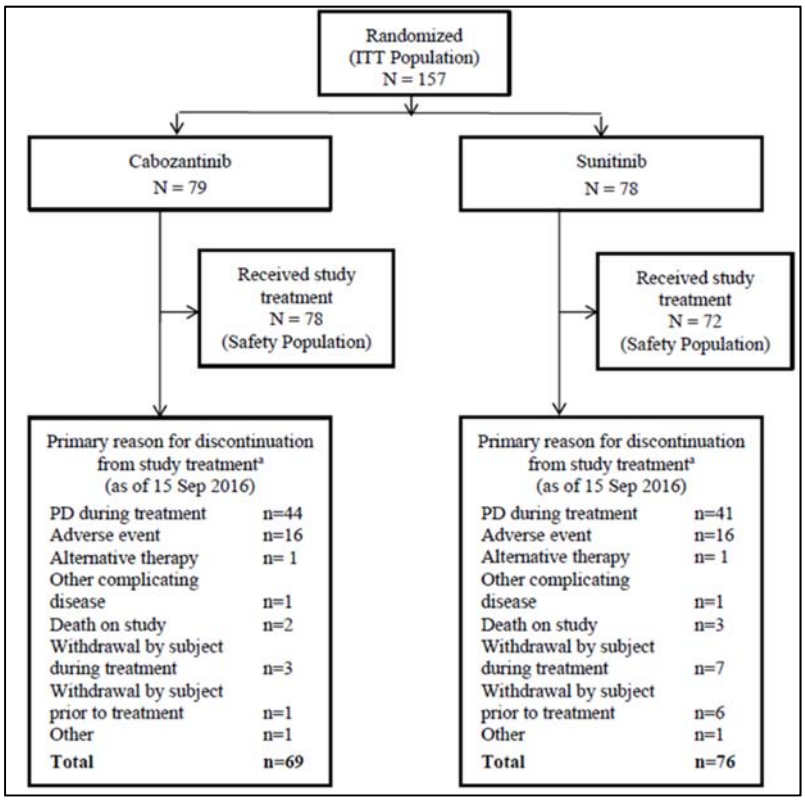
Reasons for the exclusion of studies at the full paper stage are provided (CS Figure 3) and references to these studies are listed in Appendix D1.1. Nine references are listed as 'article not obtained' in the flowchart and in response to a clarification request (question A12), the company states that these references were identified from the systematic review in the pazopanib company submission to NICE[13] (eight out of these nine references were conference abstracts). The company states that these references would either have been excluded or had been superseded by a more recent full text publication. It is our view that the non-availability of these nine references would not have biased the company's systematic review.

3.1.3 Identified studies

The CS identified one relevant published RCT, the A031203 CABOSUN trial (NCT01835158) referred to as CABOSUN for short. CABOSUN was an investigator-led phase II, open-label trial set in 77 centres in the USA from July 2013 to April 2015,

conducted by the Alliance for Clinical Trials in Oncology. The trial included adult patients (≥18 years of age) with untreated clear cell metastatic RCC, ECOG performance status of 0 to 2 and intermediate or poor risk per IMDC criteria comparing a cabozantinib treatment arm with a sunitinib treatment arm. The trial was supported by grants from the National Institutes of Health and by Exelixis (the manufacturer of cabozantinib, who provided the drug).

A CONSORT flowchart of the trial is presented in CS appendix D1.2, detailing the number of patients that discontinued/dropped out and associated reasons (see Figure 3).



Source CS appendix D.1.2 Figure 52

Figure 3 CABOSUN trial participant flow chart

Clinical effectiveness evidence is presented from the company study report (CSR)[27] and three journal publications, of which two were conference presentations.[24 28 29] The trial was used in support of the company's application for marketing authorisation, although not designed as a registration trial but used as such due to what the CS describes as "encouraging results". Due to requirements of the marketing authorisation, there are some discrepancies between the results presented in the CS and those in the trial journal publications (discussed in more detail in Section 3.1.6 of this report).

As can be seen in Table 4, the CABOSUN trial did not include any UK patients. Patients received 60 mg of cabozantinib orally once per day or 50 mg of sunitinib orally once per day (sunitinib: 4 weeks on and 2 weeks off), with a treatment cycle defined as 6 weeks for both. All patients regardless of treatment arm received full supportive care and AEs were managed through dose interruptions and dose reductions in both treatment arms.

Table 4 CABOSUN trial characteristics

<table>
<tr><th>Design, patient population and length of follow-up</th><th>Intervention</th><th>Comparator</th></tr>
<tr><td rowspan="3">Trial name: CABOSUN

Design: Phase II, open-label, multicentre RCT

Location: 77 centres in the USA

Setting: hospital and outpatient clinics

Number of participants: 157

Inclusion: Adults ≥ 18 years of age with documented RCC with some component of clear cell histology, that was advanced (defined as not amenable to curative surgery or radiation therapy) or metastatic (American Joint Committee on Cancer stage IV). Other key eligibility criteria were;
• Intermediate or poor risk by IMDC criteria
• ECOG performance status 0 to 2
• No prior systemic treatment for RCC
• No active brain metastases; patients with treated brain metastases which had been stable for at least 3 months were eligible
• Adequate organ and marrow function with no uncontrolled significant illness.

Length of follow-up:
(Randomisation July 2013 to April 2015) The median follow-up of surviving patients as of 15/09/2016 was 21.4 months.</td><td>Cabozantinib (n= 79) administered orally once per day at a dose of 60 mg</td><td>Sunitinib (n= 78) administered orally once per day at a dose of 50 mg for 4 weeks, followed by a 2-week break</td></tr>
<tr><td colspan="2">Adverse events were managed with treatment interruptions and dose reductions: cabozantinib to 40 and 20 mg, and sunitinib to 37.5 and 25 mg.

A treatment cycle was defined as 6 weeks in both study groups. Treatment duration was until disease progression, intolerance to therapy, or withdrawal of consent for treatment.</td></tr>
<tr><td colspan="2">Background therapy: all received full supportive care (including transfusions of blood and blood products, erythropoietin, antibiotics, antiemetics, and other agents) when appropriate.

Prophylactic measures were taken to prevent or reduce the severity of palmar-plantar erythrodysaesthesia syndrome (PPES; hand-foot syndrome).

Palliative radiotherapy was not permitted and concomitant use of medications that are strong inhibitors/inducers of CYP3A4 were to be avoided.</td></tr>
</table>

Source: CS Table 8 and Table 9

The CS provides a summary of the different data cut-off points used, combined with the outcome analyses and the source of the data (Table 5) and the key differences between the investigator and regulatory analyses of the trial (Table 6). As well as using different data cut-off points, the main differences between the two registration analyses appear to be the censoring rules and the use of one- or two-sided p-values (see Section 3.1.6 for more detail).

Table 5 CABOSUN data cut-off points and outcomes analysed

Date	Outcomes analysed	Source	Additional information
11 April 2016	PFS and (ORR)[a]	Choueiri 2016 et al[24]	Investigator assessment. Alliance censoring rules for progression (Missing or inadequate tumour assessments or use of systemic non-protocol anticancer therapy were not reasons for censoring) and 1-sided p-values. Event-driven analysis triggered when 123 events were observed.
15 September 2016	PFS and ORR	CSR and Choueiri et al 2017[30]	Additional analyses performed for regulatory purposes. Results in the CSR are based on assessment by an IRC and FDA-recommended censoring rules, with two-sided p-values. FDA-recommended censoring rules for PFS necessarily reduced the number of events available for analysis. To increase the number of events that would be included in the analyses, the data cut-off for radiographic endpoints in the CSR was extended to 15 September 2016 (database extract 13 January 2017 - the latest date for which OS data were available).
13 January 2017	OS (Exploratory analysis)	CSR	OS analyses were conducted with the most mature OS data available at the time.
1 July 2017	OS (Exploratory analysis)	Choueiri et al 2017[30]	Results from the updated OS analysis

Source: partly based on CS B.2.2 Table 6.
CSR, clinical study report; FDA, Food and Drug Administration; IRC, Independent Radiology Committee; ORR, objective response rate; OS, overall survival; PFS, progression-free survival.
[a] It is not clear why the objective response rate (ORR) next to progression free survival (PFS) in the first row of the CS table is bracketed.

As can be seen in Table 6, differences between investigator and regulatory analyses for CABOSUN resulted in different patient numbers for those with radiographic images and differences in the number of events recorded.

3.1.3.1 CABOSUN trial baseline characteristics

The CS states the demographic characteristics were well balanced between study groups, albeit as can be seen in Table 7, there are some exceptions. There are some differences in age range (cabozantinib 40-82 years; sunitinib 31-87 years), male sex (cabozantinib 84%; sunitinib 73%), prior nephrectomy (cabozantinib 72%; sunitinib 77%) and visceral metastases (cabozantinib 77%; sunitinib 72%). In response to a clarification request, the company confirmed that all RCC patients in the CABOSUN trial had metastatic disease (clarification question A2).

Table 6 Key differences between investigator and regulatory analyses for CABOSUN

	Original report (Choueiri 2016[24])	CSR and Choueiri 2017[30]	
Reader	**Investigator**	**Investigator**	**IRC**
No. of patients with radiographic images	157	157	156
No. of events	123	107	92
Cut-off date (PFS and ORR)	April 2016	September 2016	
Cut-off date (OS)	April 2016	January 2017 (CSR) / July 2017 (Choueiri 2017)	
Censoring rules (PFS)	Alliance	FDA guidance	
Censor for non-protocol systemic anticancer therapy	No	Yes	
Censor if event after ≥2 missing assessments	No	Yes	
Stratified analysis[a]	Yes	Yes	
P-value sided	1	2	

Source: CS B.2.2 Table 7
[a] Stratification factors: IMDC risk group (poor, intermediate) and bone metastases (yes, no).
CSR, clinical study report; FDA, Food and Drug Administration; IRC, Independent Radiology Committee; ORR, objective response rate; OS, overall survival; PFS, progression-free survival.

The ERG notes from CS Appendix D1.1 Table 11 that there were differences between the study groups in number of metastatic sites, with the percentage of patients with ≥3 sites 32% in the cabozantinib group, compared to 41% in the sunitinib group. The corresponding figures for 2 sites were 47% versus 26%, and corresponding figures for 1 site were 22% versus 33%. Thus, a greater proportion of patients in the cabozantinib arm had two or more metastatic sites (79%) than in the sunitinib arm (67%).

Clinical expert advice to the ERG suggests that these differences are not large enough to be of clinical importance. Expert clinical advice to the ERG also suggests the baseline characteristics are generally representative of patients seen in UK clinical practice apart from the proportion of patients with prior nephrectomy. This is higher than normally seen in clinical practice based on the experience of one of the experts.

3.1.3.2 Non-randomised trials

The CS for the clinical effectiveness of cabozantinib was limited to RCTs and no non-randomised studies were included in the submission.

Table 7 CABOSUN baseline patient characteristics

Characteristic, n (%)	Cabozantinib (n=79)	Sunitinib (n=78)
Age, median years (range)	63 (40-82)	64 (31-87)
Sex, male	66 (84)	57 (73)
Race		
White	70 (89)	75 (96.2)
Black	3 (4)	2 (2.6)
Asian	2 (3)	0
Other, unknown or not reported	5 (6)	1 (1)
ECOG PS		
0	36 (46)	36 (46)
1	33 (42)	32 (41)
2	10 (13)	10 (13)
IMDC risk group		
Intermediate	64 (81)	63 (81)
Poor	15 (19)	15 (19)
Bone metastases		
Yes	29 (37)	28 (36)
No	50 (63)	50 (64)
Prior nephrectomy		
Yes	57 (72)	60 (77)
No	22 (28)	18 (23)
Metastases[a]		
≥ 1 metastatic site	79 (100)	78 (100)
Visceral metastases	61 (77)	56 (72)

Source: CS Table 10

There is a small error in the CS table of baseline patient characteristics (CS Table 10), with the number of participants under race in the cabozantinib arm totalling to 80 rather than 79. It would appear that the number of Asian participants should have been one rather than two, as per the trial publication.[23]

[a], as reported by the investigator on the on-study case-report form. ECOG, Eastern Cooperative Oncology Group; IMDC, International Metastatic Renal Cell Carcinoma Database Consortium; PS, performance status.

3.1.4 Description and critique of the approach to validity assessment

The CS included a risk of bias assessment (CS Table 15 appendix D1.3) using the criteria suggested by NICE[31] for the CABOSUN and COMPARZ RCTs (details of the latter are reported in 3.1.7). Table 8 shows the company's and the ERG's quality assessment of the trial.

Table 8 Company and ERG assessment of trial quality - CABOSUN

NICE QA Criteria for RCT	CS response	ERG response
1. Was the method used to generate random allocations adequate?	Yes	Unclear risk of bias
Comments: Random stratified assignment [IMDC risk category (intermediate or poor) and presence of bone metastases (yes or no)] in a 1:1 allocation ratio using a dynamic allocation method. Dynamic allocation (DA) methods balance prognostic factors between treatment groups, which are a primarily deterministic, non-random algorithm.[32] However, DA is a family of methods, not just one, and the company does not specify which approach they used. The ERG requested details of the DA process employed (clarification question A4c), but the company did not provide any information beyond that already stated in the CS. It is therefore unclear why DA was needed given that there is already stratification, which prognostic variables were included in the DA algorithm and what part of the DA algorithm was random.		

2. Was the allocation adequately concealed?	Not clear	Unclear risk of bias
Comments: The method of allocation concealment is not reported in the trial publication, study protocol or CS.		
3. Were the groups similar at the outset of the study in terms of prognostic factors, e.g. severity of disease?	Yes	Yes (low risk of bias)
Comments: The publication and the CS state that overall, the treatment groups were balanced with respect to baseline demographic and disease characteristics. However, there were minor differences between the treatment arms, with the cabozantinib arm containing 11% more male patients, a slightly different ethnic mix (7.6% fewer white patients), 5% fewer patients who had had a prior nephrectomy, and 5% more patients with visceral metastases than the sunitinib arm. A greater proportion of patients in the cabozantinib arm had two or more metastatic sites (79%) than in the sunitinib arm (67%). Clinical expert advice to the ERG suggests that these minor differences would be unlikely to have clinical implications.		
4. 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	No (high risk of bias)
Comments: Open label trial. The CS states that a central imaging review of investigator assessments was not performed. However, a blinded central review by an IRC was undertaken retrospectively to minimise bias for the PFS and response outcomes in the company's updated analysis.		
5. Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No	Unclear risk of bias
Comments: The most frequent primary reasons for study treatment discontinuation were disease progression (cabozantinib 56%, sunitinib 53%) and AEs (20% and 21%, respectively) (clarification question A15). The company states that in general, the numbers of dropouts were considered balanced, and the ERG agrees that this is the case for withdrawal due to progression and AEs. However, there were differences between the study arms in the number of patients who did not receive the study drug (cabozantinib n=1, sunitinib n=6) and in the number of patients who withdrew consent (cabozantinib n=3, sunitinib n=7 according to Figure 52 in CS Appendix D1.2; but n=1 and n=9 respectively according to CS section B.2.13). The company states that the frequency of withdrawal by subject during treatment is considered as low (clarification question A15). We note that these withdrawals amount to 3.8% of the cabozantinib trial arm and 9.0% of the sunitinib trial arm. It is unclear whether this difference would have introduced bias, since the reasons for patients' withdrawal of consent are not reported. It should also be noted that there was an imbalance between the cabozantinib and sunitinib arms in the proportions of patients who had ≥2 missed "adequate tumour assessments" before a PFS event, and in the proportions who had no post-baseline "adequate tumour assessments". In response to a clarification request (question A7), the company states that the reasons for these differences are not available.		
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No (low risk of bias)
Comments: There are no deviations from the trial protocol with regard to outcomes.		
7. 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 (low risk of bias) Yes Yes
Comments: States that an ITT approach (defined as all patients who were randomised) was used for all but safety data (the safety analysis population was defined as patients who received ≥1 dose of study drug). In response to a clarification request on missing data, the company states that in the retrospective IRC assessment of PFS and ORR, no values were imputed for patients for whom a complete set of baseline and post-baseline radiographic images were not available and FDA censoring rules were applied (clarification question A6). The application of FDA-recommended censoring rules for PFS necessarily reduced the number of events available for analysis (CS Section B.2.2). The CS states that in the retrospective IRC assessment of PFS and ORR, no values were imputed for patients for whom a complete set of baseline and post-baseline radiographic images were not available (CS Table 11). Therefore 156 patients with radiographic images and 92 events were included in the		

retrospective analysis compared to 157 patients and 123 events in the original analysis (CS Table 7).

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 regularity 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 ≥ 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 it is an important outcome that should be included, as it generally reflects a patient's day-to-day functioning.[33] 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;[24]
- 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).[24] Results of the updated analysis are also reported in the CSR and in a conference presentation.[30]

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).

3.1.6.2 PFS (primary outcome)

The original analysis approach for PFS, as reported in the trial publication,[24] employed unblinded radiological assessments made by the trial investigators, censoring according to Alliance rules (missing or inadequate tumour assessments or use of systemic non-protocol anticancer therapy were not censored), a one-sided hypothesis test, and a data cut-off of April 2016. The company's updated analysis, as presented in the CS, required radiological assessments to be made retrospectively by a blinded IRC, censoring according to FDA rules (missing or inadequate tumour assessments or use of systemic non-protocol anticancer therapy were censored), a two-sided hypothesis test, and was based on a data cut-off of September 2016 (CS Table 7). Median follow-up for PFS in the updated analysis was 25.0 months.

The data cut-off for progression in the original analysis was event-driven, with analyses being triggered when 123 events were observed. For the updated analysis, the CS states that to increase the number of events that would be included in the analyses, the data cut-off for radiographic endpoints was extended to 15th September 2016. We note that the number of events achieved at this later cut-off (CS Table 7) was less than the 123 specified in the power calculation (see 'Sample size and power calculation' below).

3.1.6.3 OS (secondary outcome)

The original analysis of OS, as reported in the trial publication,[24] was based on a data cut-off of April 2016. The updated analysis, reported in the CSR and CS, employed a data cut-off of 13 January 2017, with a median follow-up of 28.9 months. The CS also reports an analysis of OS at the latest available data cut-off, 1 July 2017 (as reported in a conference presentation[30]) (CS Table 7). Median follow-up was not reported for this analysis.

The OS data at all the analysis time points were immature. The CS cautions that there was a notable degree of censoring around the median estimates, and confidence intervals around the hazard ratios were wide due to the relatively low number of deaths (CS section B.2.6).

The CSR states that "the study did not have a pre-specified hypothesis for the treatment effect on OS, so inference tests should be interpreted accordingly" (CSR section 11.5).

3.1.6.4 ORR (secondary outcome)

The initial and updated analysis approaches for ORR were the same as those employed for PFS (CS Table 7).

Standard statistical methods were used to compare the ORR between cabozantinib and sunitinib (CS section B.2.4). The difference in percentage ORR between groups was tested with a 2-sided Cochran-Mantel Haenszel (CMH) test with the same stratification factors as the PFS analysis. The CS clearly reports the percentage ORR with 95% confidence interval for each trial arm; the ORR treatment difference with 95% confidence interval; and the p-value for the difference from the CMH test (CS Tables 14 and 15).

In addition to the ORR, the CS reports descriptively (i.e. without statistical analysis): the numbers and percentages of patients in each trial arm with: a complete response; a confirmed partial response; stable disease; progressive disease; unevaluable or missing data; the percentage with any reduction in the target lesion; and the disease control rate (CS Tables 14 and 15). According to footnote d in CS Table 15, the CS reports the percentage with progressive disease as "progressive disease as best overall response". The company clarified that this refers to the proportion of patients whose best overall response to treatment with regard to tumour response was classified as 'progressive disease' (clarification question C1).

3.1.6.5 Analysis populations

The CS states that all efficacy analyses were carried out in the intent-to-treat (ITT) population, defined as all patients who were randomised. The safety analysis population was defined as all patients who received any treatment with cabozantinib or sunitinib. Patients were analysed according to actual treatment received (CS Table 11).

3.1.6.6 Sample size and power calculation

The null hypothesis in the initial analysis of PFS was that the HR for progression of the cabozantinib and sunitinib arms would be 1.0. The alternative 1-sided hypothesis was that the HR would be 0.67, favouring cabozantinib over sunitinib.

The CS reports that a sample size of 123 events (progressions or deaths) would provide the log-rank test with 85% power to detect a HR of 0.67 for PFS, assuming a 1-sided type I error rate of 0.12, equivalent to an increase in median PFS from 8 months in the sunitinib arm to 12 months in the cabozantinib arm (CS Table 11). Assumptions required to achieve the target of 123 events are stated in CS Table 11 (including 5.8% accrual rate over 24 months, minimum PFS follow-up 20 months, and exponential distribution of PFS).

We note that the updated analysis of PFS as reported in the CS would have been under-powered statistically compared to the initial analysis specified in the sample size calculation, since a 2-sided test has less statistical power than a 1-sided test, and 92 events occurred in

the updated analysis due to different censoring rules, which is fewer than the planned target of 123 events (CS Table 7).

3.1.6.7 Treatment of missing data

As noted above, the CS states that censoring rules for the updated analyses of PFS were applied in accordance with FDA guidance (CS Table 11); the FDA rules can be inferred from CS Table 12 and are stated explicitly by the company in their response to clarification question A6. In the retrospective IRC assessment of PFS and ORR, no values were imputed for patients for whom a complete set of baseline and post-baseline radiographic images were not available (CS Table 11).

The reasons for censoring PFS data in the retrospective IRC analyses, based on FDA rules, were: ≥2 missed analyses prior to an adequate tumour assessment (ATA); no baseline and post-baseline ATA; no event by the last ATA; no post-baseline ATA, and receipt of systemic anticancer therapy (CS Table 12). As noted above, missing or inadequate tumour assessments or use of systemic non-protocol anticancer therapy were not reasons for censoring in the initial investigator analysis approach using the Alliance censoring rules (CS Table 7).

According to CS Table 12, there were imbalances between the cabozantinib and sunitinib arms in the proportions of patients who had ≥2 missed adequate tumour assessments before a PFS event (6% versus 0%) and in the proportions who had no post-baseline adequate tumour assessments (1% versus 8%). The company explained in a clarification response that information on the reasons for these differences is not available (clarification question A7).

The CS states that there was an imbalance in the number of patients with missing data (CS section B.2.13). One patient in the cabozantinib arm and six in the sunitinib arm withdrew prior to receiving study treatment, but the reasons for these withdrawals were not known. There was also a higher incidence of missing or unevaluable data in the sunitinib arm, with six patients in the cabozantinib arm and 18 in the sunitinib arm not evaluable because they had no adequate post-baseline imaging assessments. The reasons were: cabozantinib: adverse event (n=5), withdrew consent (n=1); sunitinib: adverse event (n=6), death (n=2), disease progression (n=1), withdrew consent (n=9). We note that the numbers who withdrew consent are slightly different in CS Appendix D.1.2 Figure 52, which gives 3 and 7 in the cabozantinib and sunitinib arms respectively. The CS states that “because of the nature of these clinical events, none of these patients was likely to have experienced a response or

prolonged PFS". The CS further states that "based on their baseline characteristics (unavailable to the ERG to verify), the sunitinib patients without post-baseline imaging would not be expected to have a better prognosis than sunitinib patients who had a response recorded, and therefore it is unlikely that the radiographic endpoints were biased against sunitinib by these missing data" (CS section B.2.13). Clinical experts advising the ERG suggested that whilst this assumption may be reasonable, it is difficult to be sure (given the lack of data on the characteristics of patients with and without post-baseline imaging). Experts also commented that the 9 patients who withdrew consent in the sunitinib arm is a relatively high proportion (i.e. 11.5% of patients in the sunitinib arm) and, speculatively, might reflect their dissatisfaction with assignment to the comparator rather than to the experimental treatment. However, we note that an imbalance in the number of patients who withdrew consent was not seen in the open-label COMPARZ trial, where 6.6% and 6.7% of patients in the pazopanib and sunitinib arms withdrew consent.[34] We also note an unexplained inconsistency in the number of patients who withdrew consent in the CABOSUN trial, as reported in the CS, which differs between CS section B.2.13 (1 and 9 withdrew from each trial arm) and CS Appendix D.1.2 Figure 52 (3 and 7 withdrew).

There appears to be inconsistency in the CS regarding the number of inadequate radiographic images or tumour assessments. CS section B.2.6 states that 13 patients did not have complete data for radiographic images or tumour assessments but these do not appear to have been accounted for among the 24 patients mentioned in CS section B.2.13 (as referred to above), who did not have adequate post-baseline imaging assessments. Further, CS Table 14 suggests that the number not evaluable was 10. The company clarified that these differences are due to the timing of the assessments for ORR responses being the entire period prior to progression, while for PFS only the response at time of progression was considered. The difference in the patient numbers seen thus reflects the fact that ORR and PFS were mostly evaluated at different numbers of points (clarification question A8).

Sensitivity analyses were conducted to explore the effect of potentially informative censoring in PFS analyses based on IRC assessments; these are reported in the CSR but not in the CS. Four sensitivity analyses examined the impact of (a) discontinuation of study treatment for reasons other than radiographic progression with no non-protocol anticancer therapy (NPACT) or (b) receipt of NPACT prior to progression. The four analyses were (i) censored subjects meeting criterion (a) were classified as events in both treatment arms; (ii) censored subjects meeting criteria (a) or (b) were classified as events in both treatment arms; (iii) censored subjects in the cabozantinib arm meeting criterion (a) were classified as events in the cabozantinib arm but remained censored in the sunitinib arm; and (iv) censored subjects

in the cabozantinib arm meeting criteria (a) or (b) were classified as events in the cabozantinib arm but remained censored in the sunitinib arm (this was the most conservative analysis) (CSR Table 17).

3.1.6.8 Subgroup analyses

Subgroup analyses of PFS per IRC assessment, with censoring according to FDA rules, are mentioned briefly in CS section B.2.7 and are presented in CS Appendix E. Subgroup analyses were pre-planned, except for age, race and sex which were exploratory analyses (CS Table 9). A total of 16 HRs for cabozantinib versus sunitinib are presented in a forest plot for the following subgroups:

- The analysis stratification factors: IMDC risk category (referred to as "Heng risk factors" in CS Appendix Figure 53) (intermediate, poor); and bone metastases (yes, no);
- MET status (positive, negative, missing);
- Age, years (<65, ≥65);
- Sex (male, female);
- Race group (white, other);
- Baseline ECOG performance status (0, 1, 2).
- Bone metastases (yes, no)

For each subgroup, CS Appendix Figure 53 presents the number of events and the median PFS in in each trial arm, and the HR with 95% confidence interval. The CS, CSR and Statistical Analysis Plan do not specify whether an adjustment was made to the type I error rate to account for multiple subgroup testing. The company confirmed in a clarification response that no adjustment was made (clarification question A9). The CSR states that, for completeness, HRs (and 95% CIs) were generated regardless of the size of the subgroup (CSR section 11.4.3.9). Subgroup sizes ranged from 8 to 70 subjects in the cabozantinib arm and from 3 to 75 subjects in the sunitinib arm (CS Appendix E Figure 53).

The CSR and Statistical Analysis Plan report that further subgroup analyses of OS, ORR and PFS were conducted per investigator radiology assessment and following both Alliance and FDA censoring rules, although these analyses are not included in the CS. The company provided subgroup analysis results for OS in response to a request by the ERG (clarification question A9).

3.1.6.9 Summary of company's approach to trial statistics

Overall, the ERG agrees with the company's approach to statistical analysis, which employed standard methods. We also agree with the company's caution that the OS data at all time points are immature and should be interpreted with caution. The key limitation in the company's approach noted by the ERG is that there were unexplained imbalances between the trial arms in missing data on tumour assessments and in patient discontinuations due to withdrawal of consent, and it is unclear whether these might have introduced bias. We also note that the updated analysis of PFS is under-powered relative to the power specified in the sample size calculation. Subgroup analyses included some subgroups with small sample sizes and no adjustment was made to control the type I error rate when analysing multiple subgroups.

3.1.7 Description and critique of the company's approach to the evidence synthesis

The CS presents a narrative review of clinical effectiveness, with study characteristics and results presented in text, tables and figures. As only one RCT of cabozantinib was included in the systematic review a meta-analysis of cabozantinib trials was not possible. However, to facilitate comparison with pazopanib an indirect treatment comparison (ITC) was performed, for the outcomes of PFS and OS. The following sections describe and critique the ITC, and a tabulated critical appraisal can be found in Appendix 9.1.

3.1.7.1 ITC evidence networks

The CS reports that a total of 19 trials (n=105 records) were identified for inclusion in the ITC (CS Section B.2.9), based on criteria that included treatments for RCC within the NICE scope (cabozantinib, sunitinib, pazopanib) and treatments outside the scope (interferon alfa, interleukin-2, sorafenib, bevacizumab and interferon alfa, temsirolimus, tivozanib, placebo). The ERG notes from CS table 18 that a total of 13 RCTs (reported in 19 publications) were included in this network, which is a discrepancy with the reported 19 RCTs mentioned in the CS.

CS Figure 9 illustrates the evidence network constructed from the 13 RCTs (reproduced below in Figure 4 – it is not stated whether this network is specific to OS or PFS outcomes, or both). In this network cabozantinib is connected via sunitinib (from the CABOSUN trial), which in turn is connected to sorafenib, pazopanib, interferon alfa, and bevacizumab and interferon alfa. These treatments in turn connect to tivozanib, placebo, and temsirolimus. For some comparisons the network contains both direct and indirect evidence ("closed loops"), and for other comparisons only direct evidence is included. The CS refers to this as a potential evidence network constructed to identify additional connections between

cabozantinib and pazopanib. Hereafter the ERG refers to this as the "wider network" of 13 RCTs (i.e. containing both in-scope and out-of-scope treatments).

The CS subsequently restricted inclusion to the ITC only to studies which included the comparators relevant to the scope of the appraisal (pazopanib and sunitinib). Studies which did not include these comparators were excluded unless they provided an intermediate link. The "restricted network" included two studies: CABOSUN (comparing cabozantinib with sunitinib) and COMPARZ (comparing sunitinib with pazopanib).[34 35] The restricted evidence network therefore includes three treatments connected via a common comparator, sunitinib (CS Figure 11 reproduced below in Figure 5).

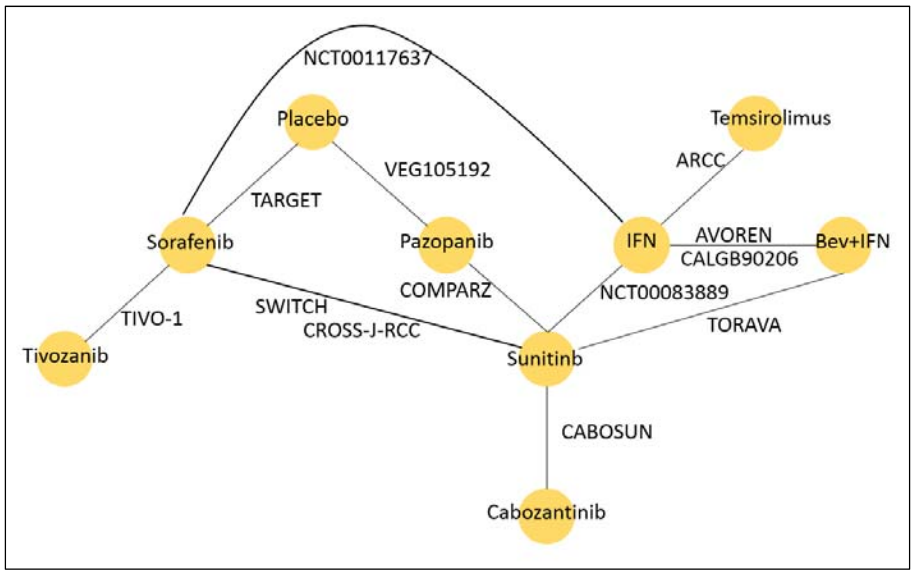


Figure 4 Wider evidence network of 13 trials (reproduced from CS Figure 9)

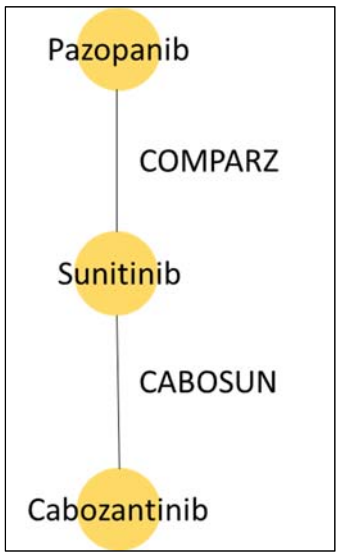


Figure 5 Restricted evidence network (reproduced from CS Figure 11)

The CS does not provide a heterogeneity assessment (statistical or clinical) of the trials in the wider network of 13 RCTs, and does not report results of any ITC based on this network. The company was requested to provide ITC results using this wider network (clarification question A26) to permit comparison of the results of the wider network with the restricted network (i.e. to check whether the results for the comparison between cabozantinib, sunitinib and pazopanib were different when a wider network containing other treatment comparisons was used). The company provided these results as survival curves, HR plots and tabulated HRs, for OS and PFS, for two analysis approaches which they had used to conduct the ITCs - Ouwens et al[36] parametric survival models and fractional polynomial models (see section 3.1.7.3 below for an explanation of these models), based on both random effects and fixed effects.

ITC feasibility assessment

The CS reports conducting a feasibility assessment for the ITC (CS section B.2.9). This assessment had two stated components: to assess whether adequate outcome data were available; and to assess whether there were differences in study and patient characteristics within and between treatment comparisons that might influence treatment effects (i.e. clinical heterogeneity). This feasibility assessment appears to have been applied only to the two trials included in the restricted ITC network (i.e. not to the wider network of 13 RCTs described above).

CS Tables 20 and 21 report the data availability assessments for PFS and OS, respectively. Hazard ratios for the ITT population (adjusted/stratified and unadjusted/unstratified) and RCC risk subgroups (intermediate risk and poor risk) are tabulated for both trials. The CS states that PFS data would be acceptable if measured either by IRC or by study investigators, with the IRC assessment considered by the company to be less likely to be biased and prioritised where possible.

Clinical heterogeneity

The CS states that the differences in distribution of RCC risk category is the variable that most affects survival (CS Appendix D). CS Table 22 compares risk category and ECOG performance status between the two included trials. The CABOSUN trial classified risk status according to the IMDC criteria (for definition of these see section 2.1.4), whilst the CS states that the COMPARZ trial used the original MSKCC criteria. However, Table 11 in Appendix D1.1 reports both IMDC and MSKCC risk classifications for the COMPARZ trial.

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 EGOG 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 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%.[37] 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).

There were slight differences between trials in the number of metastatic sites detected (≥3 sites: 32% to 41% by treatment arm in CABOSUN; 42% to 44% by treatment arm in COMPARZ). (CS Appendix Table 11). Just over a third of patients in CABOSUN had bone metastases at baseline (36% to 37% by trial arm) compared to 15% to 20% (by trial arm) of patients in COMPARZ. The CS states that patients with bone metastases have a poor prognosis and experience poorer outcomes with currently available treatments compared with patients without bone metastases. A higher percentage of patients in COMPARZ received prior nephrectomy (82% to 84% by trial arm) compared to patients in CABOSUN (72.2% to 76.9% by trial arm). Expert clinical advice to the ERG suggests this may be explained by the fact that patients with more favourable RCC risk are more likely to receive nephrectomy (COMPARZ included some patients with favourable RCC risk). Also, fewer nephrectomies tend to be performed now in practice than in the past (COMPARZ is an older trial than CABOSUN). Expert clinical advice also notes that prior nephrectomy is associated with a better treatment outcome, thus raising the potential risk of bias in the ITC results.

There were differences in ethnicity between the two trials: 92% of patients were classified as white in CABOSUN, compared to 64% white in COMPARZ (34% were described as being Asian). All patients in CABOSUN were from the USA, whereas patients in COMPARZ were from 14 countries located in North America, Europe, Australia, and Asia. The trials were comparable in terms of age (median age 61 to 64 across the trials) and reasonably similar in gender profile (male: 73.1% to 83.5% in CABOSUN; 71% to 75% in COMPARZ). The inclusion criteria of both trials required patients to have locally advanced or metastatic clear cell RCC. All patients in the CABOSUN trial had metastatic disease, whilst 98% had stage IV disease in COMPARZ.

In terms of design characteristics, the CABOSUN trial was a phase II RCT (n=157 patients, of whom 79 were randomised to cabozantinib and 78 were randomised to sunitinib), whilst COMPARZ was a larger phase IIIb non-inferiority RCT (n=1110 patients randomised, of whom 557 were randomised to pazopanib and 553 were randomised to sunitinib). The ITC is therefore unbalanced in terms of the proportions of patients randomised to the three respective treatments. The primary outcome measure in both trials was PFS.

In both trials the study treatments were administered continuously until progression of disease, the occurrence of unacceptable toxic effects, or withdrawal of consent. The dose regimen of sunitinib was identical in both trials (orally once per day at a dose of 50 mg for four weeks, followed by a two-week break).

Patient crossover was not permitted in the CABOSUN trial, and the CS states that the occurrence of crossover was not reported in the COMPARZ trial (CS Appendix D1.1 Table 10). The ERG has checked the available reports of the COMPARZ trial and can find no mention of crossover.[34 35 38 39] As reported earlier, 57 to 58% of patients in the CABOSUN trial received subsequent anticancer drug treatments following discontinuation of study treatment. In COMPARZ 55% to 56% of patients received subsequent anticancer therapy, including sunitinib in pazopanib-treated patients, and vice versa. The occurrence of subsequent anticancer treatment will affect estimates of OS in both trials.

In summary, there some similarities but also a number of differences between the two RCTs in the ITC, with the most important difference being in RCC risk status. Overall, patients in the CABOSUN trial had a poorer RCC risk status and cancer performance status than patients in the COMPARZ trial. The CS does not comment on the likely implications of this on the ITC results. The ERG considers the effect of this on the ITC results to be uncertain.

3.1.7.2 Critical appraisal of trials included in the ITC

CS appendix D provides the company's critical appraisal of the two trials included in the ITC (Figure 41 and 42 and Table 15). A brief commentary is provided in which it is stated that the trials met assessment criteria for method of randomisation, balanced trial arms at baseline, no selective reporting and use of ITT analysis. However, it is stated there was potential risk of bias due to lack of patient blinding to treatment allocation (both trials were open label), and lack of information on allocation concealment. As discussed earlier in this report (section 3.1.4), the ERG mostly agrees with the company's critical appraisal of the CABOSUN trial. The ERG also conducted an independent critical appraisal of the COMPARZ trial to compare with that of the company (see Appendix 9.2). The ERG notes that a blinded central review by an IRC was undertaken in both trials for the PFS and response outcomes (retrospectively in CABOSUN), thus the potential for detection bias is reduced for those outcomes, though performance bias is still possible.

In summary, the ERG considers the methodological quality of the two trials to be adequate overall and the overall risk of bias to be low, with the exception of bias related to lack of blinding, and bias relating to sequence generation and allocation concealment procedures which were not clearly reported. The other limitations of the CABOSUN trial need to be acknowledged, namely, the fact that it is a relatively small phase II trial with immature OS data.

3.1.7.3 Statistical ITC methods used

Three different statistical methods were used to conduct the ITC:

(1) Indirect comparison of parametric survival curves using methodology developed by Ouwens et al (2010).[36]

(2) Parametric models with fractional polynomial distributions using methodology developed by Jansen (2011).[40]

(3) A "network meta-analysis: supplementary method" comparing hazard ratios using a fixed effects model, for intermediate risk and poor risk subgroups and the ITT population.

Methods 1 and 2 were used to inform the economic model and are included in the CS due to the company's observation that the assumption of proportional hazards was violated in the CABOSUN trial for OS and PFS, and for PFS in COMPARZ (Appendix D1.1 Table 12). The ERG concurs that proportional hazards do not hold for OS in CABOSUN as the survival curves in CS Figure 6 clearly cross at around month 14. However, the PFS survival curves (CS Figure 5) appear parallel after around month three. In the COMPARZ trial the ERG concurs that proportional hazards do not appear to hold for PFS based on visual inspection of the survival curves.[34 35] However, the ERG notes that the OS survival curves in this trial appear to cross at around month 24.[35] Because of these differences in opinion the company were asked to clarify their conclusions on the proportional hazards assumptions (clarification question A3).

The company responded by supplying scaled Schoenfeld plots and log-cumulative hazard plots for OS and PFS in both trials. Plots of Schoenfeld residuals against time are a standard approach to test for the (non-)proportionality of hazards; violation of the proportional hazards assumption is indicated if the plot of Schoenfeld residuals against time shows a non-random pattern. The company state that the Schoenfeld plots show an "increasing trend followed by a decreasing trend" and that the log-cumulative plots show "roughly parallel curves". The ERG considers that proportional hazards hold for PFS but not OS in the CABOSUN trial based on inspection of the log-cumulative hazard plots. For the COMPARZ trial the reverse is apparent: proportional hazards do not appear to hold for PFS but they do for OS. Given the observation of non-proportionality of hazards in at least one of the outcomes in both trials the ERG considers use of ITC methods that accommodate time-varying HRs to be appropriate.

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 conducing an ITC.[36] 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[36] 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[36] 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 deaths for each interval and patients at risk at the beginning of the interval.[36]

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

the consistency assumption) requires covariates that act as relative treatment effect modifiers to be similar across trials. As discussed above, this assumption may not hold given the differences between the two trials in factors such as baseline RCC risk status and proportions of patients with bone metastases.

The parameter estimates for differences between treatments in scale and shape can be reported (accompanied by credible intervals), and expressed visually as HR and hazard rate plots showing treatment curves over the follow-up period.[36] The CS does not present hazard ratio or hazard rate plots, but does present fitted survival curves for the outcomes of OS and PFS for all three treatments, based on fixed effects and random effects models, for each of the five parametric survival distributions (Figure 1 to Figure 20, Appendix D1.1).

In summary, the ERG considers the Ouwens et al[36] method appropriate for implementing the ITC given the violation of the proportional hazards assumption for OS in the CABOSUN trial (notwithstanding the aforementioned caveats about clinical heterogeneity between trials). The ERG notes that this method has been used in two previous NICE appraisals, of breast cancer treatment (TA239 and TA503).[41 42]

3.1.7.5 Fractional polynomial model

The CS cites a publication by Jansen[40] as the basis of their use of fractional polynomial methodology. Jansen describes this method as an alternative to NMA of survival data in which the treatment effect is represented by a constant HR. A multi-dimensional treatment effect approach is used in which hazard functions of interventions compared in an RCT are modelled, and the difference between the parameters of these fractional polynomials within a trial are synthesized (and indirectly compared) across studies. The fractional polynomial analysis generates results which reflect the time course of the log-hazard function and as such can be expressed as log-hazard function curves and their parameters (intercept and slope). Credible interval curves can be plotted alongside the log-hazard function curves. The ERG notes that fractional polynomial-based NMAs have also been included in other NICE STAs, including appraisals of renal cell carcinoma treatments (TA463 and TA512).[14 22]

Two orders of FP model were considered for inclusion: first-order, and second order. The power level for each order can be chosen from the following set -2. -1, -0.5, 0, 0.5, 1, 2, 3. A first order model with a P= 0 would be equivalent to a Weibull model, and a first order model with P=1 would correspond to a Gompertz model. For the first order model the following powers were considered in the CS: P=-1, P=-0.5, P=0, P=0.5 and P=1. For the second order

model the following powers were considered: P1=-0.5, P2=0; P1=-1, P2=0; P1=-1, P2=-1; P1=-1, P2=0.5; P1=-1, P2=1 (see CS Table 24).

The ERG notes that only a relatively narrow range of powers (P1 and P2 in the range -1 to +1) were considered in the company's analysis. Given that none of the modelled OS curves in the CS appeared to reflect the shape of the CABOSUN KM OS curves, the company was asked if they had considered a wider range of powers (thus reflecting other functional forms) (clarification question A24). The company responded with a number of justifications for their chosen range. They stated that the joint estimation of parameters "is very delicate for every (P1, P2) model and the lack of stability of the estimation algorithms typically causes very long run times" thus they had to be strategic in their choice of which powers to test. They also cite their previous submission to NICE on cabozantinib for second line treatment of RCC and the fact that the best fitting fractional polynomial in that submission was within the same range of powers. They also state that their guiding principle was that smaller values of P1 and P2 should be preferred, implying that using higher power values would lead to over-fitting which would give curves uncharacteristic of typical PFS or OS curves. Overall, the ERG considers that the justification given by the company for the range of powers tested is reasonable.

3.1.7.6 Choice of fractional polynomial model

To select the most appropriate fractional polynomial model from the first and second order models considered, the company used the DIC to compare goodness-of-fit. The DIC is commonly used to compare the fit of Bayesian statistical models. The model with the smallest DIC is estimated to be the model that would best predict a replicate dataset which has the same structure as that currently observed.[43] The best fitting fractional polynomial model chosen for OS and PFS was the second order model P1=-1 and P2=-1 (CS Table 24), and this was used to inform the economic model (CS Table 59). The CS does not state whether any other considerations were taken into account in the choice of model, such as clinical plausibility with respect to the OS and PFS estimates generated. They comment that this model was also the best-fitting model in their previous work on cabozantinib for the treatment of second line RCC (response to clarification question A24), which the ERG assumes refers to NICE TA463.[14]

The CS presents fitted fractional polynomial survival curves (first and second order) for the outcomes of OS and PFS for all three treatments, based on fixed effects (CS Figure 21 to Figure 40, Appendix D1.1). The CS did not supply hazard ratio plots for each fractional polynomial model with credible intervals to allow visual inspection of the time-varying HRs.

These were requested from the company as well as the tabulated HRs for each interval of the follow-up period (clarification question A22). The company were also asked to provide fractional polynomial results based on a random effects model (clarification question A23). The company provided the requested data and these are described and discussed in section 3.3.7 of this report.

3.1.7.7 Bayesian statistical methods used in the Ouwens and fractional polynomials ITCs

The ERG noted that limited details of the Bayesian methods used to run both the Ouwens et al[36] and fractional polynomial models[40] are given in the CS. Details lacking included the prior probability distributions (e.g. vague, informative, non-informative, the rationale for their choice), the likelihood distribution, the number of iterations used for burn in and inferences, and the methods for assessing convergence. The company were requested to provide this information (clarification question A21).

The company reported that in the Ouwens et al[36] method non-informative priors were used for all models, with model parameters estimated using a Markov Chain Monte Carlo Gibbs algorithm in WinBUGS software. For fixed effects models, three parallel chains were run, with 50,000 iterations for burn in and a further 100,000 iterations for inferences. These were increased to 150,000 and 200,000 iterations respectively for the random effects models.

For the fractional polynomial method the choice of prior was also non-informative and a Markov Chain Monte Carlo Gibbs algorithm in WinBUGS software was also used. Three parallel chains were run with 250,000 iterations for burn in and a further 250,000 iterations for inferences. The Gelman-Rubin statistic Rhat was calculated and convergence declared when Rhat < 1.05 for both the Ouwens and fractional polynomial methods. Rhat is a standard model convergence statistic reported in WinBUGS; values close to 1.0 (i.e. <1.05) are considered indicative of convergence.[44] The company did not report whether or not they had conducted sensitivity analyses on choice of prior.

Based on the information provided the ERG considers that the methods used to implement the two ITC methods are appropriate and correspond to the methods specified in the original methodological texts.[36 40]

3.1.7.8 Summary of the ERG's appraisal of the ITC

- The company conducted ITCs to compare cabozantinib against pazopanib given the lack of head-to-head evidence for these two treatments.

- The company's ITC includes two RCTs: CABOSUN (cabozantinib versus sunitinib) and COMPARZ (sunitinib versus pazopanib). CABOSUN was a phase II RCT (n=157 patients) whilst COMPARZ was a larger phase IIIb non-inferiority RCT (n=1110 patients).
- These two trials have some similarities:
    - Treatments were administered continuously until progression of disease, the occurrence of unacceptable toxic effects, or withdrawal of consent; identical dose regimen of sunitinib were used; mean age and gender profile was similar; all patients had clear cell RCC and most patients had metastatic disease.
    - However, there are some important differences: 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; Around a third of patients in CABOSUN had bone metastases (a key prognostic factor in RCC) at baseline compared to 18% of patients in COMPARZ; a greater proportion of COMPARZ patients were classified as having the highest cancer performance status. The impact of these differences on the results of the ITC are not discussed in the CS. The ERG considers that they may under-estimate the relative effectiveness of cabozantinib versus pazopanib.
- The ERG considers the methodological quality of the two trials to be adequate overall and the overall risk of bias to be low, though there is risk of bias relating to blinding due to the open-label nature of the trials. The CABOSUN trial has some further limitations (i.e. phase II trial, relatively small sample size; immature OS data).
- Due to the observation that proportional hazards do not hold for all survival outcomes in both trials, the CS used ITC methods that do not assume proportionality in hazards. These were the ITC of parametric survival curves using methodology developed by Ouwens et al,[36] and the use of parametric models with fractional polynomial distributions using methodology developed by Jansen et al.[40] The Ouwens et al method provides survival estimates for a family of parametric distributions (Weibull, log-logistic, log-normal, Gompertz, exponential). The fractional polynomial method provides survival estimates for first order and second order models from a set of powers (five models for each order, 10 models in total). Both of these methods provide survival effect estimates that are used in the company's economic model.
- The Ouwens and fractional polynomial methods appear to have been implemented adequately in accordance with the original publications,[36 40] and the ERG considers

that both are suitable for use for the indirect comparison of treatments in this appraisal. However, the results of both methods may be biased by the aforementioned differences between the two trials in RCC risk factors and other variables.

- The results of the ITC based on these methods are described later in this report (section 3.3.7 and their suitability for use in the economic model to inform cost-effectiveness estimates are discussed in section 4.3.4).

3.2 Summary statement of company's approach to evidence synthesis

The ERG's assessment of the company's approach to the evidence synthesis is summarised in Table 9.

Table 9 Quality assessment (CRD criteria) of CS review

CRD Quality Item	ERG response
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	Yes. Eligibility criteria are tabulated (CS Table 4) and are generally appropriate, but with the following minor inconsistencies: • The tabulated inclusion criteria for interventions and comparators are broader than those finally applied to identify eligible studies. Final eligibility criteria were stated as "only treatments in the NICE scope of the appraisal" (footnote in CS Table 4) and "only publications related to cabozantinib, sunitinib and pazopanib were included in the final selection" (CS section B.2.1). Response rates are listed in the company's decision problem but are not specified as an outcome in the inclusion criteria.
2. Is there evidence of a substantial effort to search for all relevant research?	Yes. Systematic literature searches were based on a search conducted by the manufacturer of pazopanib (1980-2009) which the company updated to June 2017 and widened to include cabozantinib and tivozanib (CS section B.2.1). The overall search was comprehensive and wider than the NICE scope, although the company did not systematically search specific conferences. The ERG ran updated searches to March 2018 and did not identify any further relevant studies.
3. Is the validity of included studies adequately assessed?	Yes. The company assessed the risk of bias in the CABOSUN trial, as well as in the COMPARZ trial that was included in the company's ITC analysis (Table 15 in CS Appendix D).
4. Is sufficient detail of the individual studies presented?	Partly. The study methods (CS Tables 6-9 and 11), baseline characteristics of the participants (CS Table 10), and participant flow (Figure 52 in CS Appendix D) are clearly reported for the CABOSUN trial. Baseline characteristics of the COMPARZ trial included in the company's ITC analysis are also clearly reported (Table 11 in CS Appendix D), but limited detail on the COMPARZ trial methods is provided (Table 10 in CS Appendix D) and patient flow is not reported.
5. Are the primary studies summarised appropriately?	Yes. Results from the CABOSUN trial are clearly summarised for all clinical effectiveness outcomes (CS section B.2.6). Results from the COMPARZ trial are

	summarised in CS Tables 20 for those outcomes relevant to the ITC (PFS and OS).

The company's evidence synthesis is generally well structured and clearly reported. The company's search for clinical effectiveness studies identified a broader range of interventions and comparators than those specified in the NICE scope. The company subsequently restricted the intervention and comparators at the eligibility screening step to be consistent with the scope. Overall, the company's evidence synthesis is consistent with their decision problem and the NICE scope, with the exception that HRQoL, which is an outcome specified in the scope, was not reported in the CABOSUN trial. HRQoL is therefore not included in the company's decision problem and not reported in the clinical effectiveness synthesis in the CS; a separate systematic review of utility studies was conducted to inform the company's economic analysis (CS section B.3.4).

3.3 Summary of submitted evidence

In the following sub-sections we summarise the results of the CABOSUN trial, based on data reported in the CS for the IRC assessment to determine progression and response and the FDA recommended censoring rules, for the most recent data cut-off date available. These are based on data in the CSR[27] and a 2017 European Society for Medical Oncology (EMSO) conference poster.[30] We do not present results from the earlier trial journal publication,[24] as these are based on an earlier data cut-off date (April 2016); are based on investigator rather than IRC assessment; use non-FDA censoring rules (and were not used in the company's regulatory submission); and are not used in the economic model.

3.3.1 Summary of results for progression free survival (PFS)

Figure 6 shows the Kaplan-Meier PFS curves, based on the September 2016 data cut-off.[27] [30] At a median follow-up of 25 months, median PFS was 8.6 months (95% CI 6.8, 14.0) for cabozantinib and 5.3 months (95% CI 3.0, 8.2) for sunitinib (p=0.0008). The median difference was 3.3 months. The HR, stratified by IMDC risk category and bone metastases, was 0.48 (95% CI 0.31, 0.74). As can be seen from Figure 6, the survival curves appear parallel from month three onwards, implying proportional hazards. The majority of events recorded were for documented disease progression: 40 (51%) in the cabozantinib group; 43 (55%) in the sunitinib group. The remaining patients were censored: 36 (46%) in the cabozantinib group; 29 (37%) in the sunitinib group (CS Table 12). PFS at 12 months (% event free) was 43.1 and 21.1 in the cabozantinib and sunitinib groups, respectively.

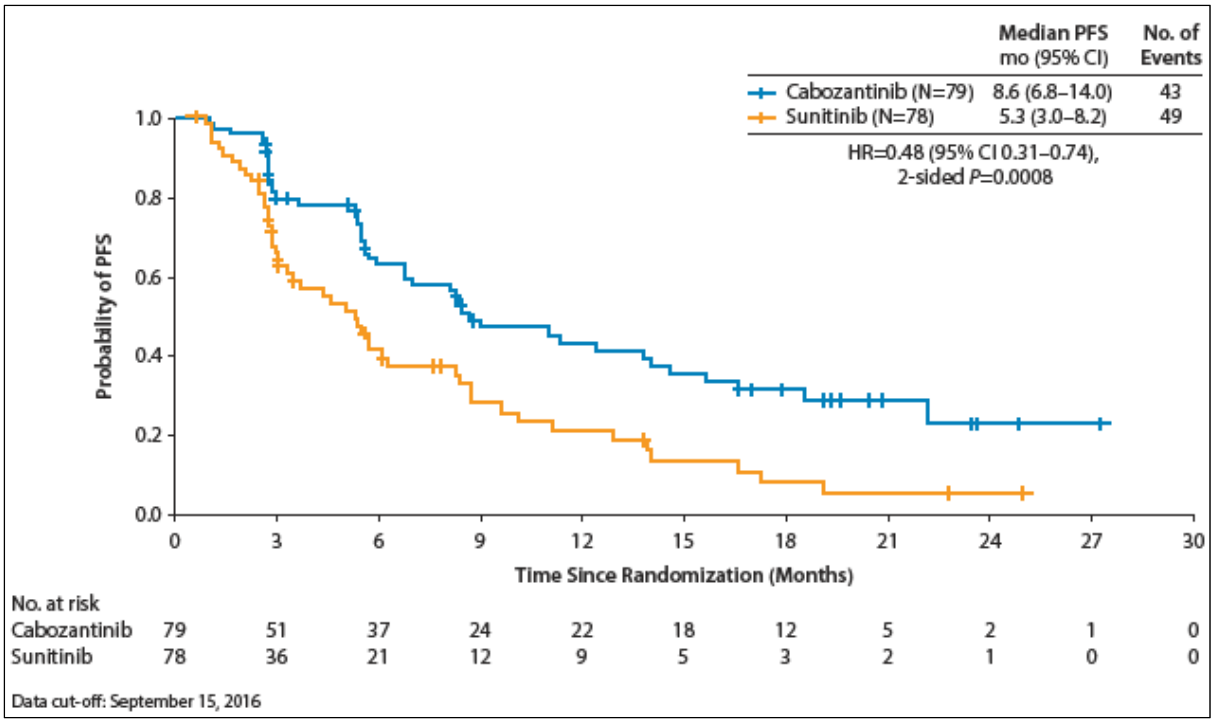


Figure 6 Kaplan-Meier PFS curves (IRC, ITT population. Reproduced from CS Figure 5)

As mentioned earlier, PFS was the primary outcome of the CABOSUN trial. However, the ERG notes that this would have been under-powered statistically, since a 2-sided test, as used in the IRC-based analysis conducted for the submission to the regulator (and also used in the CS) has less statistical power than a 1-sided test (used in the original trial analysis), and the 92 events is fewer than the planned target of 123 events (see section 3.1.6 of this report for more information on the statistical procedures used in the trial).

3.3.2 Summary of results for overall survival (OS)

Overall survival was a secondary outcome of the CABOSUN trial. Figure 7 shows the Kaplan-Meier OS curves, based on the January 2017 data cut-off. At a median follow-up of 28.9 months, the median OS was 30.3 months (95% CI 14.6, not estimable) in the cabozantinib arm versus 21.0 months (95% CI 16.3, 27.0) in the sunitinib arm. The median difference was 9.3 months. The CS notes that the data were immature at this data cut-off and there was a notable degree of censoring around the median estimates (censoring due to no event as of the cut-off date – 52% and 42% of patients in the cabozantinib and sunitinib groups, respectively). Thus the OS data should be interpreted with caution. The HR, stratified by IMDC risk category and bone metastases, was 0.74 (95% CI 0.47, 1.14) p=0.1700. As can be seen from Figure 7, the survival curves cross at around month 14 before crossing again and then separating at around month 21 for the rest of the follow-up period. Proportional hazards do not therefore appear to hold.

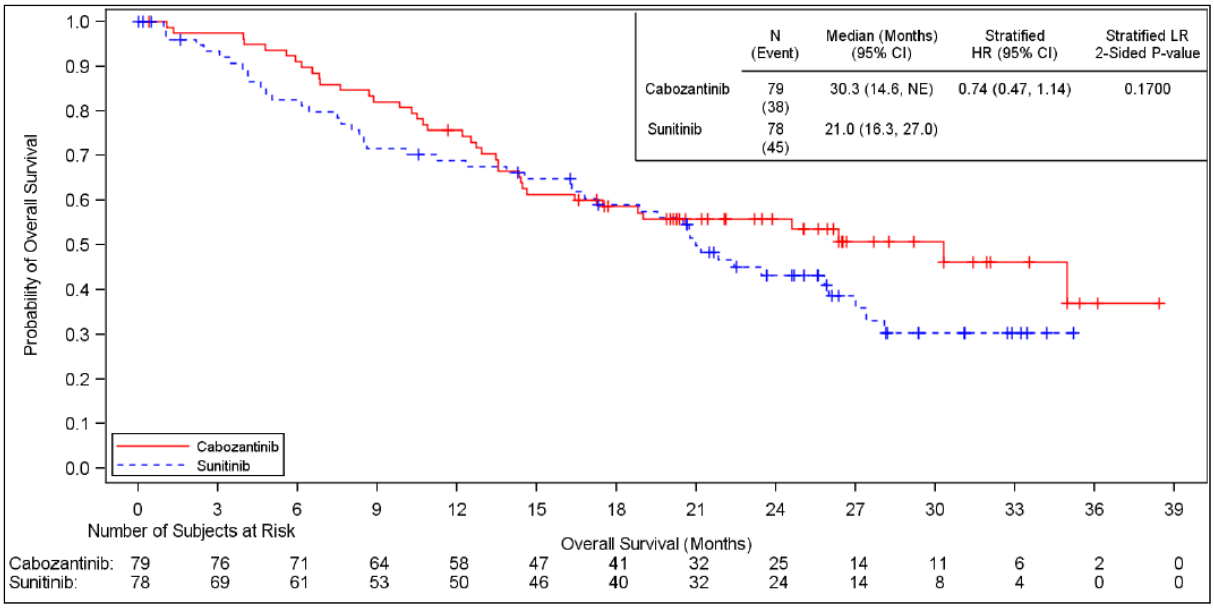


Figure 7 Kaplan-Meier plot of OS (13th January 2017 data cut-off, ITT population. Reproduced from CS figure 6)

The percentage of patients event-free at 30 months was 50.7% and 30.3% for cabozantinib and sunitinib, respectively (CS Table 13). The CS states that these OS data are used to inform the economic model.

Figure 8 illustrates an updated OS analysis, at the data cut-off of 1st July 2017 (thus around six months after the above OS analysis; median follow-up not reported), presented at the EMSO conference.[30] As can be seen, the median difference in OS between the treatments is 5.5 months, favouring cabozantinib. However, the confidence intervals around the OS estimates are wide and the confidence interval around the HR crosses 1, indicating a non-statistically significant difference. Data from this cut-off do not appear to have been used in the economic model, and it is not stated in the CS why data from the earlier OS data cut-off (January 2017) were used in preference.

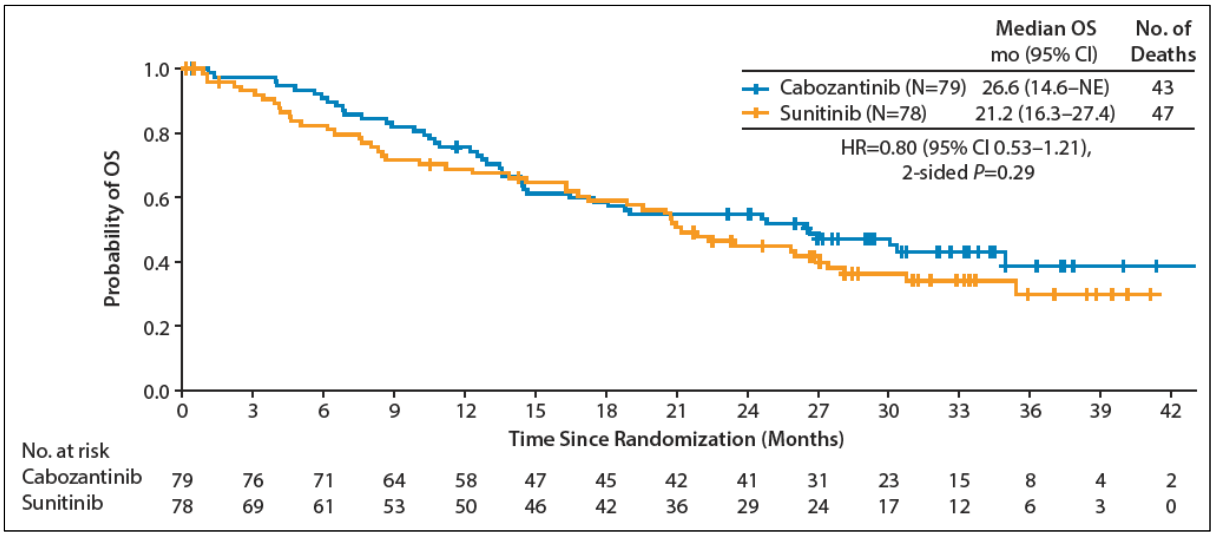


Figure 8 Kaplan-Meier plot of OS (July 2017 data cut-off, ITT population. Reproduced from CS figure 7)

The ERG notes that the OS estimates presented will have been influenced by subsequent anticancer treatments that trial participants received on discontinuation of the study treatment (systemic non-radiation anticancer therapy was received by 57%-58% of patients) (see section 3.3.4 below). The CS does not discuss the impact of these treatments on the OS estimates, or make any adjustments to the OS estimates in the economic model. The impact of subsequent anti-cancer treatments on OS is therefore uncertain. (NB. The ERG adjusts the costs to reflect different assumptions about subsequent anticancer treatments in a scenario analysis – see section 4.4).

In summary, the ERG urges caution in the interpretation of the OS results from this study as the data are immature, the survival curves cross each other indicating non-proportional hazards, the study was not statistically powered for OS, and the uncertain influence of subsequent anticancer treatments received by a large proportion of patients in both study groups.

3.3.3 Summary of results for tumour response

CS Table 14 presents tumour response data based on IRC assessment in the ITT population (data cut-off September 2016). As stated earlier, this outcome is not used to inform the economic model. The objective response rate (ORR) was 20% (95% CI 12.0%, 30.8%) in the cabozantinib arm, compared to 9% (95% CI 3.7%, 17.6%) in the sunitinib arm, classed as a 'confirmed partial response'. The difference between groups in ORR was 11.3% (95% CI, 0.4 22.2%; p=0.0406). There were no confirmed complete responders in either study group. The proportion of patients with stable disease was higher in the

cabozantinib group (n=43; 54%) than the sunitinib group (n=30; 38%). The proportion of patients with progressive disease was lower in the cabozantinib group (n=14; 18%) than in the sunitinib group (n=23; 29%). The CS also reports the disease control rate, defined as complete response + partial response + stable disease at 75% and 47% in the cabozantinib and sunitinib groups, respectively. In actuality, this rate only reflects partial response and stable disease as there were no complete responses in the trial.

CS Table 15 reports tumour response results for the three sets analyses (Investigator-assessed, Alliance censoring rules, April 2016 cut-off; Investigator-assessed, FDA censoring rules, September 2016 cut-off ; IRC-assessed, FDA censoring rules, September 2016 cut-off). This permits side-by-side comparison of the results at different cut-offs/tumour assessment/censoring rules. We have not reproduced this table here, but in summary we note that the ORR for both study groups is lower (and the between group difference is smaller) for the IRC assessment using FDA censoring rules (i.e. the data used in the company's regulatory assessment).

3.3.4 Subsequent anti-cancer treatment

Although crossover was not permitted in the trial, patients could receive subsequent non-protocol treatments upon discontinuation of the study treatment (e.g. on disease progression). Whilst not an outcome in the scope of the appraisal, the CS reports details of subsequent treatments given (CS Table 16 and Appendix L). The proportion of patients receiving any systemic non-radiation anti-cancer was similar: 45 (57%) and 45 (58%) in the cabozantinib and sunitinib groups, respectively. The median time to first systemic non-radiation anti-cancer therapy was 196 (range 56, 877) days and 147 (range 4, 725 days) in the cabozantinib and sunitinib arms, respectively. Just under half of all patients received a VEGFR-targeted TKI drug as a subsequent treatment (axitinib, pazopanib, sunitinib, cabozantinib, or sorafenib). The ERG notes that sunitinib and sorafenib have not been recommended by NICE for second-line RCC treatment. Around 14% of patients overall received an anti-PD-1/PD-L 1 targeting agent as subsequent therapy, including nivolumab. Other systematic therapies used as subsequent treatment included temsirolimus, everolimus and bevacizumab.

The company were asked to clarify the number of patients who received each subsequent line of systemic anticancer therapy (e.g. second, third, fourth line etc) (clarification question A19). The company clarified that only first non-protocol treatments and concomitant medications were captured in the case report forms. Thus it appears that the data provided on subsequent treatments refer to second-line treatment only. However, in contradiction, the

ERG notes that Appendix L Table 50 states to include all reported subsequent anti-cancer treatments (including "first subsequent treatment and any further treatments reported").

3.3.5 Summary of Health related quality of life (HRQoL)

As stated earlier, HRQoL was not measured in the CABOSUN trial. For details of the company's HRQoL utility estimates in the economic model see section 4.3.5 of this report.

3.3.6 Sub-group analyses results

CS Appendix E Figure 53 provides a forest plot showing pre-specified subgroup analyses for the outcome of PFS as determined by IRC assessment (see section 3.1.6 earlier in this report for details of the statistical procedures used in the subgroup analyses). The CS comments that there was a consistently favourable effect for cabozantinib compared with sunitinib in larger subgroups (≥ 20 patients). There was a favourable effect for cabozantinib compared to sunitinib in the following subgroups: age (<65 years ≥ 65 years); sex (male); race (white); baseline ECOG status (0); bone metastases (yes/no); RCC risk factors (intermediate/poor); and MET status (positive). Confidence intervals were wide and included 1 for some the smaller subgroups (e.g. race group 'other'). Of note, the PFS HR was more favourable for the poor RCC risk subgroup (0.31, 95% CI 0.11, 0.92) than the intermediate risk group (0.52, 95% CI 0.32, 0.82), though this is based on a very small sample of patients (15 poor risk patients in each group).

The CS did not present subgroup analyses for the outcome of OS, but supplied them on request (clarification question A9) in a table, with no commentary or interpretation. The results appear to be based on the January 2017 data cut-off. Overall, the results were consistent with the overall population analysis results, with OS more favourable in the cabozantinib group than the sunitinib group. However, in most subgroups the confidence intervals included 1 (as in the overall population analysis). Tests for treatment by subgroup interaction yielded non statistically significant p values except for MET status.

The CS does not present subgroup analyses for the outcome of tumour response. The ERG identified a conference abstract presented at the American Society of Clinical Oncology (ASCO) in February 2018 which reported subgroup analyses of the CABOSUN trial for the PFS and ORR outcomes.[26] The PFS results are the same as those reported in the CS (summarised above). Odds ratios for ORR are given for the following subgroups: IMDC risk group, bone metastases, age, sex, baseline ECOG and MET status. No confidence intervals around the odds ratios are given, or any other descriptive statistical information. The data

show odds ratios greater than 1 for all subgroups, and the abstract states that odds ratios favours cabozantinib over sunitinib. The ERG interprets this as a higher odds of achieving a confirmed partial response with cabozantinib (as was the case in the overall study population – see section 3.3.3 above).

The ERG urges caution in the interpretation of subgroup analyses as they are not statistically powered to detect a difference between treatments, and some of the subgroups are quite small leading to uncertainty in effects. In particular, the OS subgroup results require caution for the aforementioned limitations of immature data, non-proportional hazards and uncertain influence of subsequent anticancer treatments. To reiterate, the scope of this appraisal does not specify any relevant subgroups for assessment.

3.3.7 Indirect treatment comparison results

3.3.7.1 ITC results: comparison of parametric survival curves

The CS reports the results of the ITC as fitted survival curves for the outcomes of OS and PFS for all three treatments (cabozantinib, sunitinib, pazopanib), based on fixed effects and random effects, for each of the five parametric distributions generated by the Ouwens et al method.[36] For each of the analyses cabozantinib had a higher survival estimate than sunitinib or pazopanib.

It is not practical to show all of the graphs here, but to illustrate, Figure 9 below shows the PFS fitted curves based on the log-normal model which was selected by the company as the most appropriate model to inform the economic model (CS Table 33). Figure 10 below reports the OS fitted curves based on the exponential model as this was selected by the company as the most appropriate model to inform the economic model (CS Table 33). The sunitinib and pazopanib curves were similar to each other in shape and position, indicating similar effectiveness, as was the case in all of the other fitted parametric survival models (CS appendix D1.1).

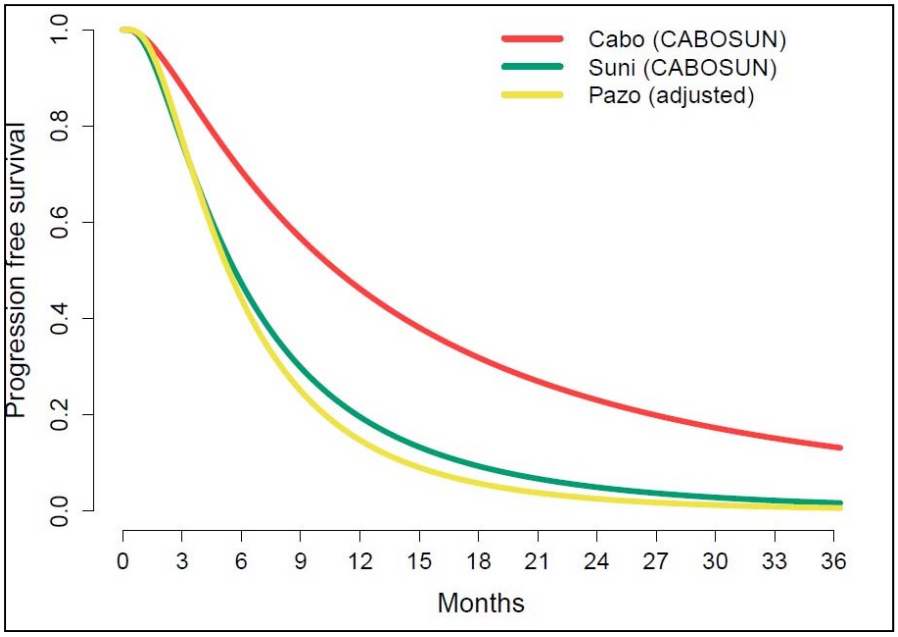


Figure 9 PFS ITC results, Ouwens model, log-normal distribution, fixed effect (reproduced from CS Appendix D1.1 Figure 14).

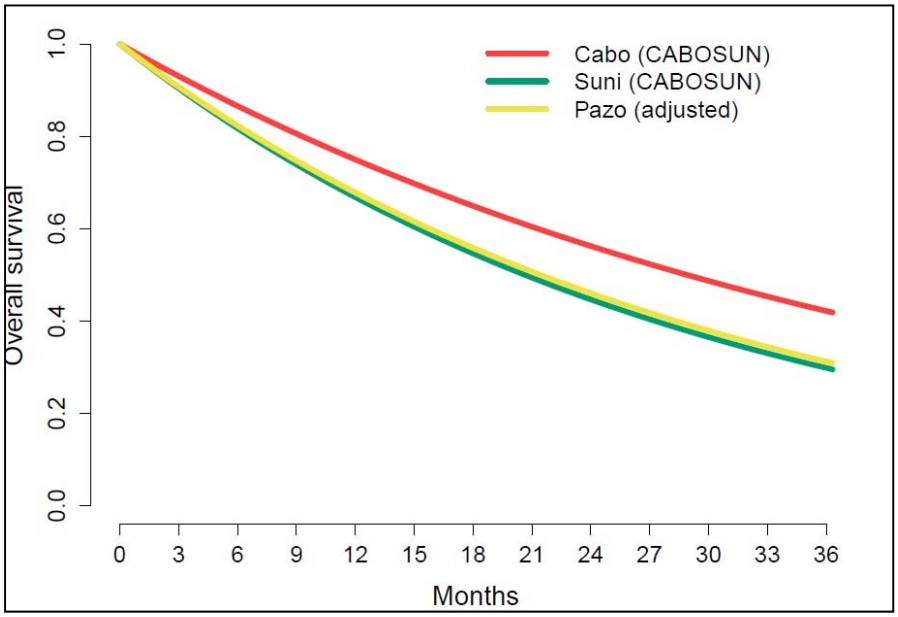


Figure 10 OS ITC results, Ouwens model, exponential distribution, fixed effect (reproduced from CS Appendix D1.1 Figure 1).

3.3.7.2 ITC results: fractional polynomials

The CS presents fitted fractional polynomial survival curves for the outcomes of OS and PFS for all three treatments, based on fixed effects for first and second order models (CS Figure 21 to Figure 40, Appendix D1.1). The CS did not supply hazard ratio plots with credible intervals for each fractional polynomial model to allow visual inspection of the time-varying HR curves. These were requested from the company as were the tabulated HRs for each

time interval of the follow-up period (clarification question A22). These were provided by the company with the time period split into monthly intervals.

It is not practical to show all of the graphs here, but for illustration, Figure 11 shows the PFS hazard ratio plot for the company's best-fitting fractional polynomial model (second order P1=-1 and P2=-1) which informed the economic model (the tabulated HRs for these plots are reported in Table 7 and Table 8 of the company's clarification question A22 response). As can be seen:

- The HR for pazopanib peaks at month four [redacted] and declines slightly during the rest of the follow-up period. The HR for sunitinib peaks at month six [redacted] and declines slightly during the remainder of the follow-up period.
- The credible intervals increase over the follow-up period, with the upper bound increasing to include 1 after month 19 for pazopanib, and after month 11 for sunitinib.
- The ERG notes that the time-varying PFS HRs for cabozantinib versus sunitinib generated by this fractional polynomial model ITC compare broadly with the constant HR reported in the CABOSUN trial (0.48 (95% CI 0.31, 0.74), though there is greater uncertainty in the fractional polynomial model as evident from the wide credible intervals which include 1 for a large proportion of the follow-up period.

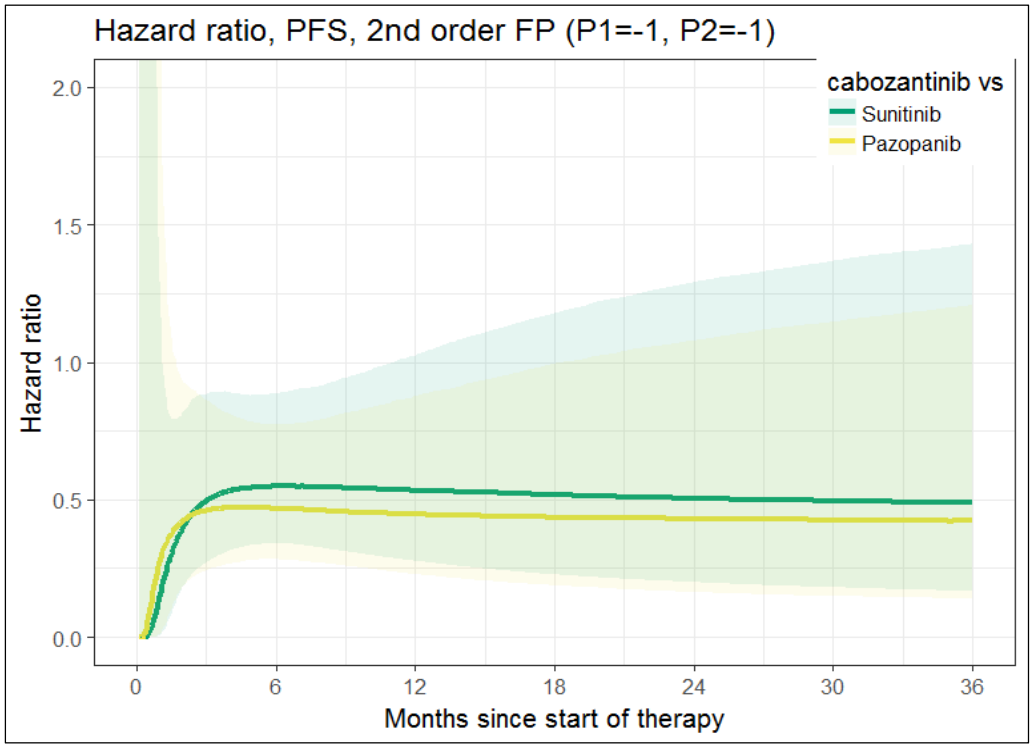


Figure 11 Hazard ratio plot, PFS; fractional polynomial 2nd order (p1=-1, p2=-1), fixed effect (reproduced from company clarification question response A22 CS figure 28)

Figure 12 shows the OS hazard ratio plot for the company's best-fitting fractional polynomial model (second order P1=-1 and P2=-1) used to inform the economic model (the tabulated HRs for these plots are reported in Table 3 and Table 4 of the company's clarification question response).

- The HR for pazopanib starts to peak at month nine, and declines slightly after month 19 [redacted] The HR for sunitinib begins to plateau at month 13 and peaks at month 30 where it remains for the rest of the follow-up period [redacted].
- The credible intervals widen during the course of the follow-up period, and include 1 at all time points.
- The ERG notes that the time-varying OS HRs for cabozantinib versus sunitinib generated by the fractional polynomials ITC compare broadly with the constant OS HR reported in the CABOSUN trial (0.80 (95% CI 0.53, 1.21), though there was greater uncertainty in the fractional polynomial model as evident from the wide credible intervals.

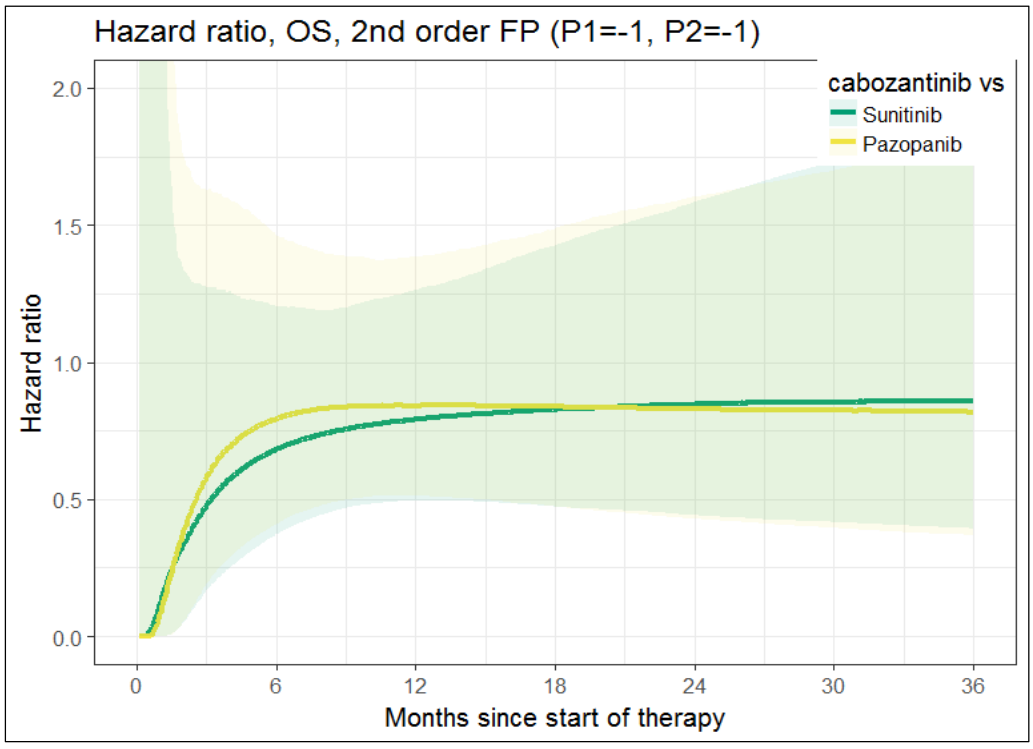


Figure 12 Hazard ratio plot, OS; fractional polynomial 2nd order (p1=-1, p2=-1), fixed effect (reproduced from company clarification question response A22 Figure 18)

The ERG has reviewed the results of the other fractional polynomial models (as supplied in response to clarification question A22, Figures 11 to 30). Our general observation is that, across the different models, the time-varying HR curves for cabozantinib versus sunitinib

and cabozantinib versus pazopanib have a similar shape to each other. Cabozantinib is of superior effectiveness when compared with both sunitinib and with pazopanib, with little difference between the results of each pairwise comparison.

Appendix 9.4 of this report provides additional ITC results:

- A summary of the results of the other (i.e. the non-best fitting) fixed effect fractional polynomial models.
- A comparison of the results of random effects and the fixed effect fractional polynomial models.
- A comparison of the results from the ITC using the wider evidence network with the restricted evidence network.

3.3.8 Summary of adverse events

CS section B.2.10 summarises adverse reactions recorded in the CABOSUN trial. Table 10 below summarises the incidence of adverse events. As mentioned earlier (section 3.1.6), adverse events were assessed in the safety analysis population, defined as all patients who received any treatment with cabozantinib or sunitinib. The safety population comprises 78/79 (99%) of patients randomised to the cabozantinib group, and 72/78 (92%) of patients randomised to sunitinib. Thus, there was a slight imbalance in the size of the study groups in this population. Adverse events were described as solicited (expected per the protocol and presence/absence and severity solicited at baseline and for each treatment cycle), and unsolicited (other adverse events not expected). The CS states that the safety data reported are taken from the CSR[27] and may differ from the trial journal publication[24] due to regulatory reporting requirements. The ERG notes that the safety data do indeed differ between these two publications, and the data in the CS (i.e. based on the CSR) should therefore be considered definitive. These data are summarised below.

The duration of treatment exposure was longer in the cabozantinib arm compared with the sunitinib arm (median: 6.5 months versus 3.1 months). Dose reductions were reported to be frequent with both treatments: (46% of cabozantinib patients; 35% of sunitinib patients) as were dose interruptions (73% and 71% respectively).

The percentage of patients with at least one treatment-related adverse events was similar between the two study groups (95%-97%). Grade 3 or 4 adverse events were reported in a similar percentage of patients in the study groups (60%-63%), as were serious adverse events (49%-51%) and treated-related serious adverse events (36%). Discontinuations of

study drug due to adverse events was also similar between study groups (21%-22%). The percentage of patients dying up to 30 days after last dose of study treatment was higher in the sunitinib group compared to the cabozantinib group (11% versus 5.1%, respectively), as was the case for death > 30 days after last dose of study treatment (49% versus 44%, respectively).

Table 10 Summary of AE incidence (safety population) (reproduced from CS Table 25)

	Cabozantinib N = 78 n (%)	**Sunitinib N = 72 n (%)**
AE	75 (96)	71 (99)
Related AE	74 (95)	70 (97)
Worst AE, grade 3 or 4	53 (68)	47 (65)
Worst related AE, grade 3 or 4	47 (60)	45 (63)
Grade 5 AE up to 30 days after last dose of study treatment[a]	3 (3.8)	6 (8.3)
Grade 5 AE > 30 days after last dose of study treatment	1 (1.3)	3 (4.2)
Related grade 5 AE at any time	2 (2.6)	4 (5.6)
Serious AE	38 (49)	37 (51)
Related serious AE[b]	28 (36)	26 (36)
Deaths	38 (49)	43 (60)
Death up to 30 days after last dose of study treatment	4 (5.1)	8 (11)
Death > 30 days after last dose of study treatment	34 (44)	35 (49)
Discontinuation of study due to AE[c]	21%	22%

[a] Grade 5 AEs were not reported for 3 subjects (1 cabozantinib, 2 sunitinib) who died < 30 days after the last dose of study treatment; [b] grade 1 or 2 SAEs that did not entail hospitalisation ≥ 24 h were not recorded in the clinical database; [c] based on patient disposition, not excluding events of disease progression, only % reported. 'Unsolicited' grade 1 and 2 events not related to study treatment were not collected.
AE, adverse event.

CS Table 26 reports the percentage of patients experiencing specific treatment-related adverse events. The incidence of specific events varied between the study groups. Common adverse events (of any grade) in the cabozantinib arm were diarrhoea (72%), fatigue (62%), aspartate aminotransferase increased (60%), hypertension (56%), alanine aminotransferase increased (54%), decreased appetite (45%) and palmar-plantar erythrodysaesthesia syndrome (42%). Of these, all except decreased appetite was a solicited adverse event. In the sunitinib group common adverse events (of any grade) included fatigue (67%), platelet count decreased (58%), diarrhoea (49%), anaemia (44%) hypertension (38%), nausea (36%) and neutrophil count decreased (35%). Of these, all except anaemia and nausea were a solicited adverse event.

Common grade 3/4 adverse events in the cabozantinib arm included hypertension (22%), diarrhoea (9%), hypophosphataemia (9%), palmar-plantar erythrodysaesthesia syndrome (7.7%), fatigue (5.1%), decreased appetite (5.1%), and stomatitis (5.1%). In the sunitinib arm common grade 3/4 adverse events included hypertension (18%), fatigue (17%), platelet count decreased (11%), diarrhoea (8.3%), and hypophosphataemia (6.9%).

Similar specific adverse events (of any grade, and grade 3/4) were common in both treatment groups, though the percentage of patients experiencing them varied between the groups.

4 COST EFFECTIVENESS

4.1 Overview of company's economic evaluation

The company's submission to NICE includes:

i) A review of published economic evaluations of cabozantinib compared with sunitinib and pazopanib for patients with untreated locally advanced or metastatic RCC.

ii) A report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of cabozantinib is compared with sunitinib and pazopanib for treatment-naïve patients with advanced RCC.

4.2 Company's review of published economic evaluations

The company conducted a systematic search of the literature to identify economic evaluations with cabozantinib or its comparators sunitinib, pazopanib in untreated advanced RCC. Details of the review methods are reported in CS Appendix G. It included cost-effectiveness studies of selected first line treatment options (sunitinib, pazopanib, interferon-alfa, interleukin-2, bevacizumab + interferon-alfa, temsirolimus, sorafenib and tivozanib) for patients with advanced/metastatic, previously untreated RCC. The search was not restricted by timeframe, language (other than English, German, French, Spanish and Italian were excluded) or countries (other than European countries, Australia, Canada were excluded).

The inclusion criteria state that full-text publications, conference abstracts and reports were included while letters, editorials, notes, and historical articles were excluded. The search identified 804 papers, which were assessed against predefined inclusion/exclusion criteria (Appendix G, Table 17). One cost-effectiveness study was excluded due to language barriers (Czech Republic). A total of 35 studies were excluded due to a focus on different countries. Table 21 (CS, Appendix G) presents the references excluded due to country.

Of the 23 studies included in CS (Table 22, CS Appendix G), 9 were critically appraised using the Drummond and Jefferson checklist (1996). Of the remaining studies, 9 were only available as conference abstracts or posters and 5 were technology appraisals published by technology assessment agencies (4 of them by NICE). Summary results of the critical appraisal are presented in Tables 23 and 24 (CS, Appendix G). Table 25 shows the studies that were not assessed with reasons (CS, Appendix G).

Table 29 (CS B.3.1) summarises the methods and results of seven studies that were conducted from an English, Welsh or British perspective. The company concluded that as

none of these studies included cabozantinib, they are not directly relevant to this appraisal. The ERG agrees with this conclusion.

4.3 Company's submitted economic evaluation

4.3.1 NICE reference case

Table 11 shows that the company's economic evaluation adheres to NICE's reference case requirements.

Table 11 NICE reference case requirements

NICE reference case requirements:	Included in submission	Comment
Decision problem: As per NICE scope	Yes	Although PFS and OS curves from ITC also include patients with favourable risk status.
Comparator: As listed in NICE scope	Yes	
Perspective on costs: NHS and PSS	Yes	
Evidence on resource use and costs: Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes	
Perspective on outcomes: All direct health effects, whether for patients or, when relevant, carers	Yes	
Type of economic evaluation: Cost utility analysis with fully incremental analysis	Yes	
Synthesis of evidence on outcomes: Based on a systematic review	Yes	
Time horizon: Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes	20 years in base case (10 in scenario analysis)
Measuring and valuing health effects: Health effect should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life.	Yes	HRQoL not collected in CABOSUN. EQ-5D estimates from published sources used
Source of data for measurement of health-related quality of life: Reported directly by patients and/or carers.	Yes	
Source of preference data: Representative sample of the UK population	Yes	
Equity considerations: An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	Yes	
Discount rate: 3.5% per year for costs and health effects	Yes	

4.3.2 Modelled decision problem

The model broadly reflects the decision problem in the scope, but with some uncertainties.

Population: The model uses a cohort with an initial age (62.8 years) and gender mix (78% male) similar to that in the CABOSUN and COMPARZ populations (Table 12). The ERG has been advised that in practice, patients starting first-line treatment for advanced RCC are often older than trial participants. We explore the impact of age on cost-effectiveness through scenario analysis to assess the applicability of the results.

Table 12 Population characteristics in the model and comparative statistics

Baseline characteristics	Model	CABOSUN[24]	COMPARZ[34]	IMDC database cohort[7]
Age (years)	62.8	Median 63	Median 61/62	55% 60+
Male	78.3%	78%	73%	74%
Favourable risk	Not explicit	0%	25%	18%
Intermediate risk		81%	55%	52%
Poor risk		19%	18%	30%

The distribution by IMDC risk group is not specified in the model but is set implicitly by the sources of effectiveness evidence. As discussed in section 3.1.7.1 above, there is an important question over how well the ITC model reflects the scope population because of the inclusion of favourable risk patients in COMPARZ. We consider the implications of this potential source of bias in relation to the choice of PFS and OS effectiveness parameters for the model.

Subgroups: The CS does not present cost-effectiveness for any patient subgroups (CS B.3.9). This is in accordance with the scope, and the company notes that CABOSUN showed consistent results across a range of subgroups (CS Appendix E). The ERG agrees that investigation of cost-effectiveness for subgroups is not warranted given available evidence, but we urge caution over interpretation of the subgroup analyses of trial data as these are not powered to detect a difference (section 3.3.6 above).

Intervention and comparators: The model compares the cost-effectiveness of first-line cabozantinib in comparison with sunitinib and pazopanib, as specified in the scope (CS B.3.2). NICE guidance recommending tivozanib in this indication[22] was published after finalisation of the scope and submission of the CS, so is not included as a comparator in the company model. We do not consider this further.

Outcomes: The model reflects the outcomes specified in the scope. Quality of life data was not collected in CABOSUN, so utilities for health states and adverse events are based on published sources for patients receiving other treatments (CS B.1.1). We discuss the appropriateness of utility sources in section 4.3.5 below.

4.3.3 Model structure and assumptions

The model structure is described in CS B.3.2 and illustrated in Figure 12, reproduced in Figure 13 below. It is a health state transition (Markov type) model, containing three mutually-exclusive health states: progression free (PF); progressed disease (PD) and death. Patients start in the PF state, following initiation of one of the included treatments at first-line: cabozantinib, sunitinib or pazopanib. At disease progression, patients transition to the PD state, which is considered irreversible, so patients cannot return from PD to PFS. Patients in PF and PD states may die from cancer or other causes.

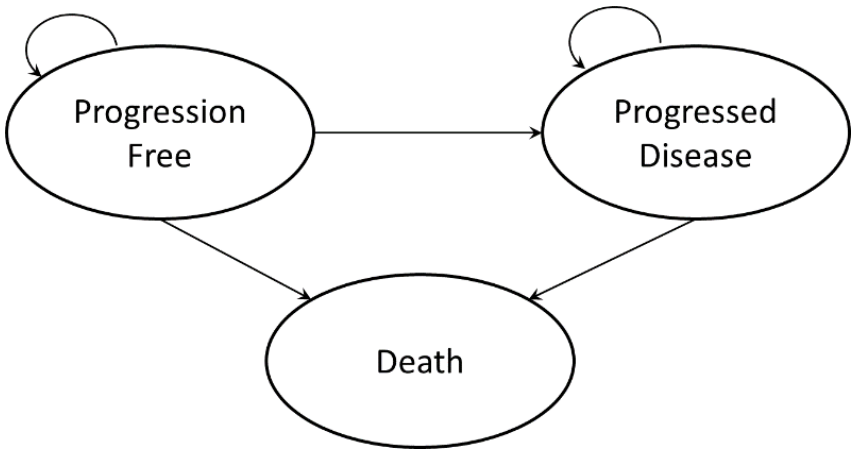


Figure 13 Structure of economic model (reproduced from CS B.3.2 Figure 12)

Alongside the health state transition model, proportions of patients on targeted treatments are estimated as illustrated in Figure 14. Patients enter the PF state on first-line treatment but may stop at any time due to adverse effects or when their disease progresses. Most patients then progress to treatment with one of 10 drugs included in the company's base case after a fixed period of waiting (8 weeks). The duration of second-line treatment is defined for each drug, after which patients are assumed to receive supportive care until death.

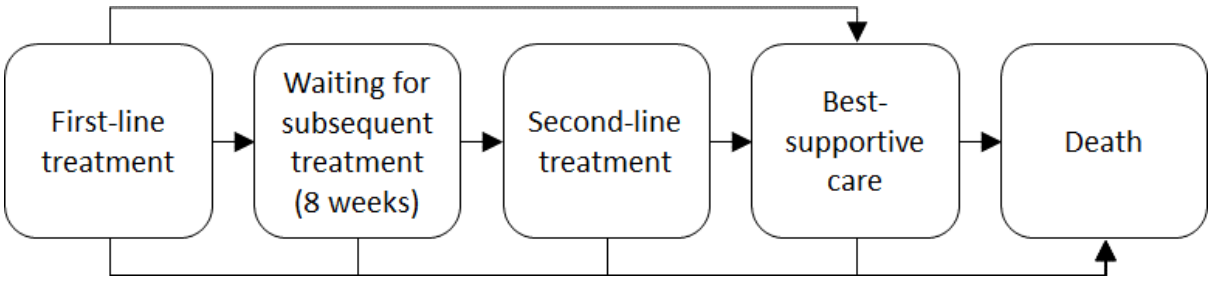


Figure 14 Treatment transition model

The distribution of the cohort between the health states at each time point is estimated using a **partitioned survival approach**, based on PFS and OS curves for the treatment arm:

- **Death**: The proportion of patients alive at each time point is taken from the OS curve. Hence the proportions of the cohort who have died are calculated.
- **PF**: The proportion of patients who are progression free is the minimum of the PFS curve and the OS curve at each time point.
- **PD**: The proportion of patients in the PD state is calculated as the residual (if any) of the cohort who are not dead and not progression free.

Similarly, the distribution of the cohort by treatment status is defined by a Time to Discontinuation (TTD) curve for first-line treatment, a waiting time of 8 weeks between first and second line and fixed treatment durations for the second-line drugs, in addition to PFS and OS curves:

- **First-line treatment**: Calculated from the minimum of the PFS and TTD curves.
- **Waiting for second-line**: The proportion of patients that start waiting in each cycle is calculated based on the proportions who are alive and end first-line treatment. The number of patients waiting is then accumulated over 8 weeks.
- **Second-line treatment**: The proportion of patients emerging alive from the waiting period is calculated and distributed between the 10 active second-line treatments and best supportive care. The time that patients spend on second-line treatment is defined by fixed treatment durations, again adjusting for any deaths within this time.
- **Best supportive care**: Patients who survive the period of second-line treatment enter the best supportive care state, where they remain until they die or the end of the time horizon.

The three-state PF/PD/death model is commonly used in cancer economic evaluations and has been used for previous NICE appraisals for untreated advanced RCC. There is some

controversy over the partitioned survival approach, however, because further assumptions are needed to estimate transition probabilities from survival curves. In this case, the submitted model assumes that the mortality rate is the same pre and post disease progression. This is unlikely but does not affect QALY or health state estimates, which are calculated from the numbers of patients in the three health states at each time point, rather than from the numbers of transitions. The model also assumes the same mortality rate for patients before and after discontinuation of first-line treatment. This does not affect the estimated duration or cost of first-line treatment, which is based on the fitted TTD curve (or PFS if lower). However, it does affect the modelled cost of second-line treatments. If the mortality rate is higher after first-line treatment than before, which seems likely, the model will tend to over-estimate the average duration and cost of second-line therapy.

Other key features and assumptions of the model are listed below:

- **Cycle length**: 1 week, with half cycle correction.
- **Time horizon**: 20 years in base case (with 10 years in scenario analysis).
- **Duration of treatment effects**: based on extrapolation of PFS and OS curves fitted to trial data, assuming persistence of effects over the time horizon.
- **Adverse events**: For each first-line treatment, grade 3 or 4 Treatment Emergent Adverse Events (TEAEs) with an incidence of 5% or more are included in the model. There is no explicit modelling of adverse events related to subsequent treatments.
- **Utility and QALY calculations**: Utility weights for the PF and PD health states are based on published estimates, assumed independent of treatment. Additional disutilities are applied to reflect included TEAEs for first-line treatments – applied as a one-off QALY loss in the first cycle. QALYs are also adjusted for the gender mix and age of the cohort.[45]
- **Health resource use and costs**: The model estimates costs associated with: acquisition and administration of first-line and subsequent treatments, with adjustment for dose intensity and wastage when appropriate; monitoring and disease management in PF and PD states; treatment of included TEAEs for first-line treatments; and end of life care applied in the last cycle before death.
- **Discounting**: 3.5% per year for costs and QALYs
- **Uncertainty**: the model includes macros to conduct: deterministic sensitivity analysis (DSA) with results presented in a tornado diagram; scenario analyses varying selected model assumptions; and probabilistic sensitivity analysis (PSA) producing a cost-effectiveness scatterplot and cost-effectiveness acceptability curve.

The ERG believes that the model structure and partitioned survival approach is appropriate, although we do have some concerns over the following issues:

- It is appropriate to estimate costs and health effects over the patients' whole lifetimes, so we do not disagree per se with the company's use of a 20-year time horizon. Other RCC appraisals have adopted a more conservative time horizon of only 10 years.[12 13 22] In the company's base case model, a relatively small proportion of the modelled cohort survive to 10 or 20 years: about 2% and 0.03% respectively with sunitinib based on CABOSUN survival data. However, we do question the extrapolation of OS and PFS curves from limited trial follow-up over 20 years. This entails strong assumptions about persistence of treatment effects, which may not be realistic. We investigate the impact of the time horizon and different assumptions about persistence of treatment effects in the ERG analysis.
- The model does not include an adjustment for age-related increase in mortality in the general population, as the it relies entirely on the projected OS curves. Given the high rate of mortality for people with advanced RCC, this might not affect results, but 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.
- The assumption of equal mortality rates before and after discontinuation of first-line treatment might lead to over-estimation of second-line treatment costs. We investigate the importance of this potential bias through sensitivity analysis on the duration of second-line treatments.

4.3.4 Treatment effectiveness and extrapolation

To apply the partitioned survival model described above, OS, PFS and TTD curves are required for cabozantinib and comparators, extrapolated over the 20-year time horizon. The company's approach to estimating these curves is described in section B.3.3 of the CS. They present two sets of base case results:

1. **Direct comparison** (cabozantinib vs. sunitinib)
   This analysis is based on patient-level data from the CABOSUN trial, with OS, PFS and TTD curves separately fitted for cabozantinib and sunitinib arms using six families of survival functions: exponential, Weibull, Gompertz, lognormal, loglogistic and generalised gamma. For their direct base case, the company chose an exponential distribution for OS and lognormal distributions for PFS and TTD.
2. **Indirect comparison** (cabozantinib vs. sunitinib and pazopanib)
   ITC meta-analyses were conducted to fit PFS and OS curves to regenerated KM

data from CABOSUN and COMPARZ, as discussed in section 3.1.7 above. Two methods were used:

a. ITC parametric curves, fixed and random effect models for five survival functions: exponential, Weibull, Gompertz, lognormal, loglogistic (Ouwens et al. method).[36] The generalised gamma distribution was not implemented due to the lack of the incomplete gamma function in WinBUGS software. The company reports that treatment was tested as a covariate, but the model only includes curves that were fitted separately for cabozantinib and sunitinib.
b. ITC fractional polynomial (FP) curves, fixed effect, with five first-order and five second order functions (Jansen method).[40] For their ITC base case, the company chose the second-order FP model with P1=P2=-1 for PFS and OS.

As TTD KM plots are not available for COMPARZ, the company uses the CABOSUN lognormal curves for cabozantinib and sunitinib, and assumes the latter would also apply to pazopanib.

We describe and critique the company's choice of OS, PFS and TTD curves below. Further critique and explanation for the ERG's preferred approach is given in section 4.4.1.

4.3.4.1 Overall Survival (OS)

OS direct comparison

The company's preferred model for OS is the exponential, with Weibull and Gompertz tested in scenario analysis. They state that this decision was based on the Survival Model Selection Process (SMEEP) from NICE Decision Support Unit Technical Support Document 14.[46]

- **Proportional hazards (PH):** The company states that PH does not hold for OS in CABOSUN. This is apparent from the KM plots (CS B.2.6 Figures 6 and 7) which cross, and the Schoenfeld and log cumulative hazard plots support this conclusion (response to clarification question A3).
- **Goodness-of-fit (AIC/AICC/BIC):** Statistical measures of fit for OS are shown in CS Tables 34 and 35. There is inconsistency between treatments. For cabozantinib, the lognormal has the best BIC, followed closely by exponential and loglogistic. But for sunitinib, Gompertz has the best BIC followed by exponential and Weibull. The company uses the exponential for both arms in their base case, arguing that this has a reasonable fit for both cabozantinib and sunitinib.

- **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.[47] 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 does not state which dataset was used, but the January 2017 KM plot is reproduced in the economic chapter (CS B.3.3 Figure 13) and KM data provided by the company in response to a clarification question also relates to this earlier 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

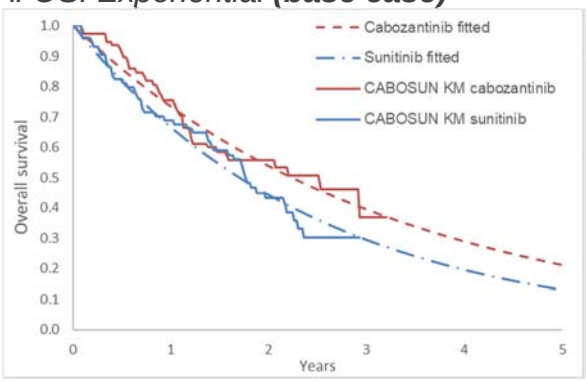
Figure 16 shows the ITC parametric and best-fitting FP survival curves in relation to the CABOSUN KM curves. We omit the COMPARZ KM curves from these graphs for clarity; but note that they are similar to the CABOSUN KM curve for cabozantinib and lie above the CABOSUN KM curve for sunitinib. This reflects 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:

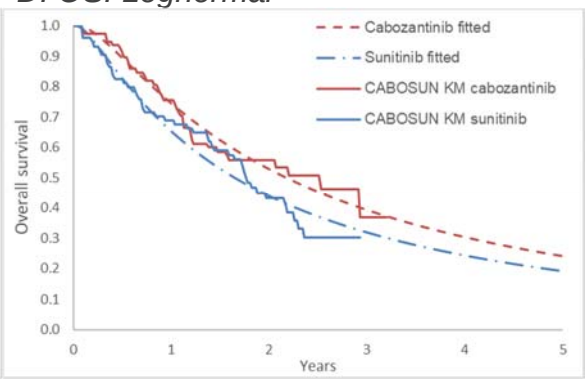
- **Proportional hazards (PH):** The company conclude that the proportional hazards assumption holds for OS in COMPARZ but not in CABOSUN. The ERG agrees with this conclusion.
- **Goodness-of-fit (DIC):** Measures of statistical fit are shown in CS Tables 23 and 24 (B.2.9). The company state that they selected the second-order FP with P1=P2=-1 because it had the best DIC statistic. They note that the first-order FPs have higher DIC statistics than second-order models (clarification question A25), so are not used in scenario analysis.
- **Plausibility of extrapolation:** The company state that two of the second-order FP models (P1=-1, P2=0.5) and (P1=-1, P2=1) are not recommended because they have "unreasonably flat tails". We note that this can also be said of the lognormal and loglogistic parametric models.

There is uncertainty over the robustness of the ITC results due to differences in the trial populations. The CABOSUN OS KM curves are also noisy, reflecting the small sample size and relative immaturity of the data. This makes it difficult to assess the fit and extrapolation of the 20 ITC curves included in the model. We consider that the RE exponential and FP P1=P2=-1 OS curves are both reasonable, with no clear reason to choose between them. The Weibull appears similar but with rather lower estimates of long-term survival with standard treatment. Conversely, the lognormal and loglogistic curves and two FP curves that the company includes in scenario analysis give high estimates of long-term survival, which we consider unrealistic. We therefore focus on the RE exponential, FP P1=P2=-1 and Weibull functions for OS in ERG additional analysis. We also consider the likely effect of using the most recent OS data from CABOSUN (July 2017 cut-off) to model cost-effectiveness.

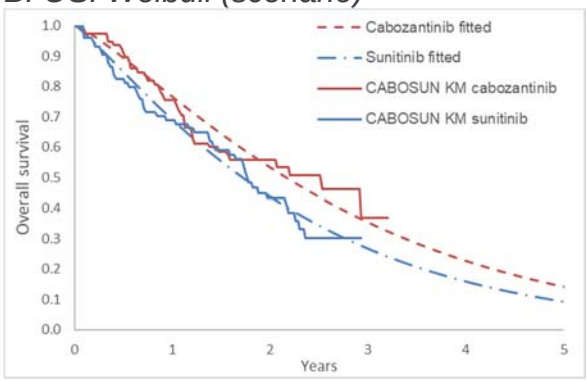
A: OS: Exponential ***(base case)***



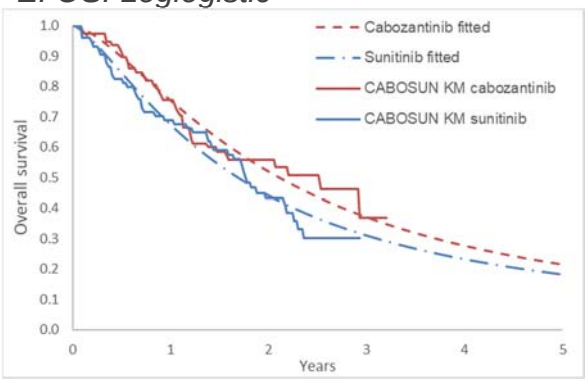
D: OS: Lognormal



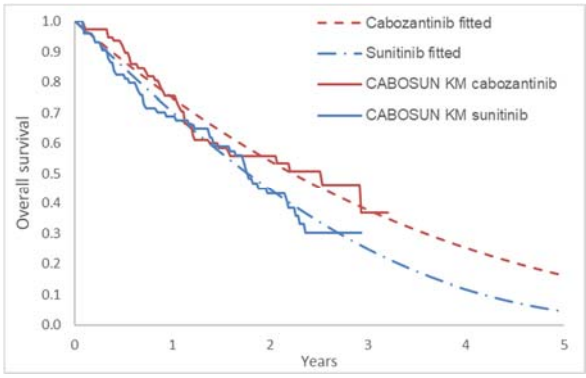
B: OS: Weibull (scenario)



E: OS: Loglogistic



C: OS: Gompertz (scenario)



F: OS: Generalised gamma

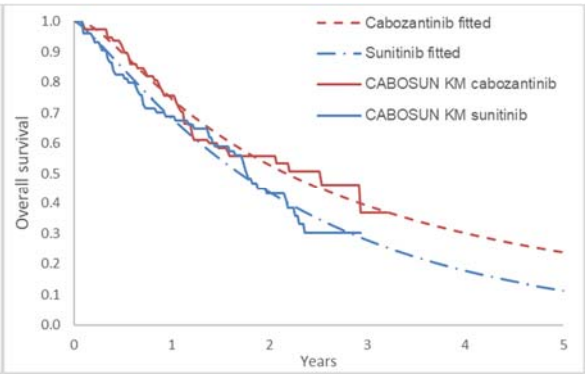
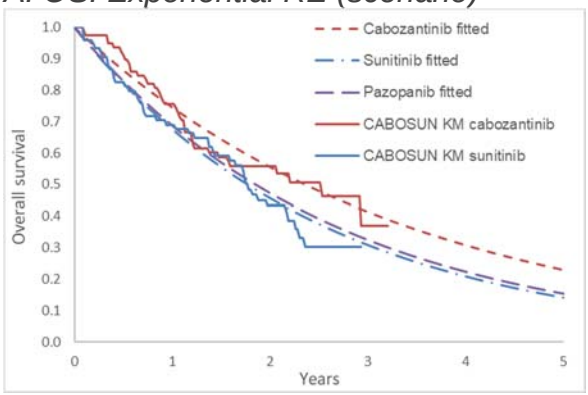


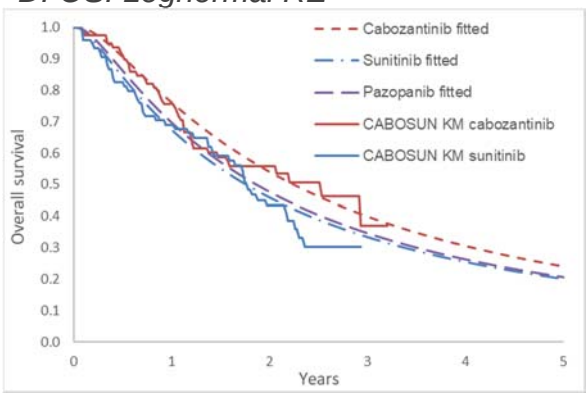
Figure 15 OS curves - fitted to CABOSUN data (direct comparison)

Source: Figures generated by ERG from company model and KM data used to fit models (Jan 2017 data cut-off)

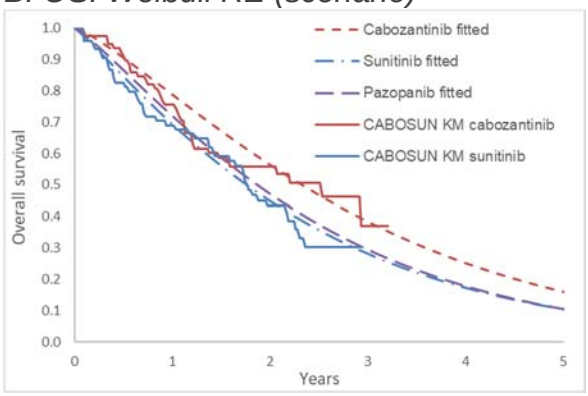
A: OS: Exponential RE (scenario)



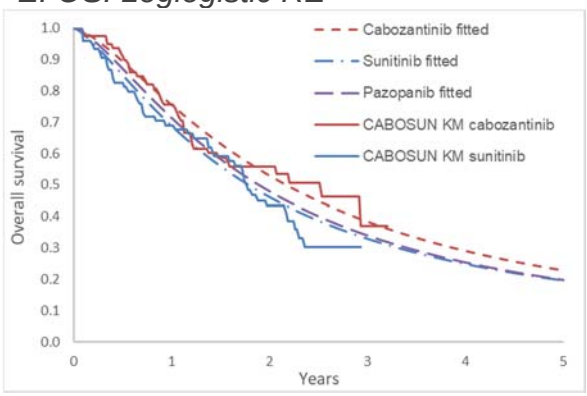
D: OS: Lognormal RE



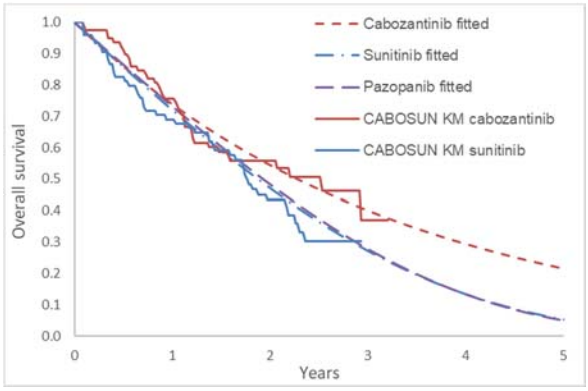
B: OS: Weibull RE (scenario)



E: OS: Loglogistic RE



C: OS: Gompertz RE (scenario)



F: OS: FP P1=P2=-1 ***(base case)***

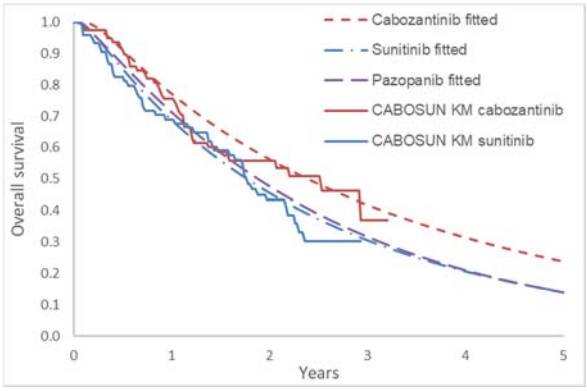


Figure 16 OS curves – ITC models fitted to CABOSUN AND COMPARZ

Source: Figures generated by ERG from company model and KM data used to fit models (Jan 2017 data cut-off). RE= Random effects

Table 13 Summary statistics for OS curves

	Model fit statistics [a]		Median OS (months)			5-year OS (%)		
			Cabo	Suni	Pazo	Cabo	Suni	Pazo
Data sources								
CABOSUN (Jan 17)	-		30.3	21.0	-			
CABOSUN (July 17)	-		26.6	21.2	-			
COMPARZ [b]	-		-	29.3	28.4			
IMDC [c]	-		-	22.3	22.6	-	21%	24%
Tivozanib STA [d]	-		-	27.5	29.2			
Fitted models: direct comparison (CABOSUN)								
Exponential	358.3	398.6	27.1	20.6	-	21%	13%	-
Weibull	360.7	401.9	26.1	20.8	-	14%	9%	-
Gompertz	362.5	394.5	26.9	21.4	-	16%	4%	-
Lognormal	358.3	403.0	26.3	20.0	-	24%	19%	-
Loglogistic	358.7	403.0	25.5	20.4	-	22%	18%	-
Gamma	362.6	406.2	26.3	20.5	-	24%	11%	-
Fitted models: ITC parametric random effects (CABOSUN & COMPARZ)								
RE Exponential	1768.9		28.6	21.4	22.4	23%	14%	15%
RE Weibull	1757.2		28.2	21.4	22.7	16%	11%	11%
RE Gompertz	1775.0		27.6	22.8	23.4	22%	6%	5%
RE Lognormal	1713.2		27.2	21.4	22.6	24%	20%	21%
RE Loglogistic	1733.4		26.3	21.7	22.9	23%	20%	20%
Fitted models: ITC fractional polynomials (CABOSUN & COMPARZ)								
FP P=-1	1722.8		27.8	21.6	22.6	18%	12%	12%
FP P=-0.5	1739.5		27.8	21.7	22.8	17%	11%	11%
FP P=0	1757.7		27.5	22.0	23.3	17%	10%	10%
FP P=0.5	1769.0		27.7	22.2	23.5	18%	8%	8%
FP P=1	1773.0		28.0	22.3	23.6	21%	7%	6%
FP P1=-0.5, P2=0	1716.5		28.4	20.8	22.9	23%	19%	19%
FP P1=-1, P2=0	1713.9		28.6	21.2	23.0	23%	17%	16%
FP P1=-1, P2=-1	1711.9		29.0	21.5	22.8	24%	14%	14%
FP P1=-1, P2=0.5	1716.2		29.0	21.3	23.1	28%	15%	15%
FP P1=-1, P2=1	1718.3		29.3	21.5	23.1	34%	14%	13%

a As reported in CS Tables 23, 24, 34 and 35: direct comparison Bayesian Information Criterion (BIC) for cabozantinib / sunitinib; parametric ITC: Deviance Information Criterion (DIC) for fixed/random effects models; and ITC FPs: Deviance Information Criterion (DIC) for first/second order models

b Motzer et al. 2013 analysis of COMPARZ trial data.[34]

c Ruiz-Morales et al. analysis of 7438 patients with metastatic RCC treated at first line with sunitinib (n=6519) or pazopanib (n=919)[47]

d ERG preferred results from Tivozanib STA (TA512)[22]

4.3.4.2 Progression free survival (PFS)

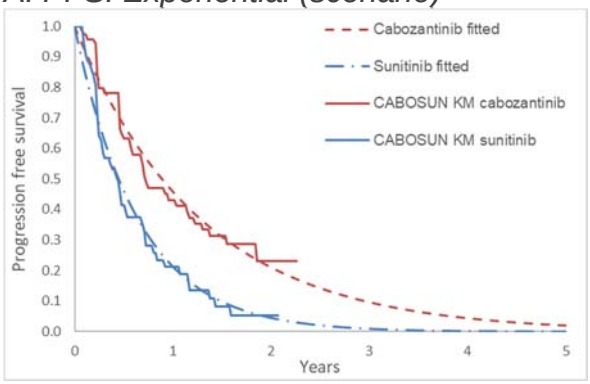
PFS direct comparison

The KM plot of PFS from CABOSUN is shown in CS B.3.3 Figure 14. We show selected graphs comparing the company's fitted curves with the CABOSUN KM plots in Figure 17 and summary statistics in Table 14 below. The KM plots from COMPARZ are higher than the KM for the sunitinib arm in CABOSUN. This is expected given the lower risk status of the COMPARZ population.

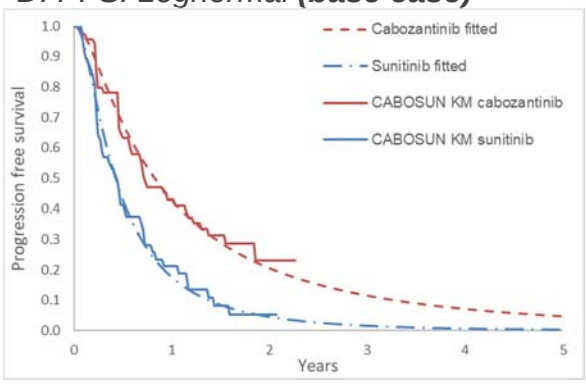
The company use separately fitted lognormal distributions for PFS in their direct base case analysis, with exponential, Weibull and Gompertz distributions in scenario analysis. They state that they made this choice based on the following considerations:

- **Proportional hazards (PH):** The company argues that the PH assumption is not appropriate for PFS in CABOSUN (CS Appendix D, Table 12). However, the ERG considers that this conclusion was not supported by the proportionality test or by the Schoenfeld and log-cumulative hazard plots (see section 3.1.7.3 above).
- **Goodness-of-fit (AIC/AICC/BIC):** Evidence of the fit of the parametric curves to trial data is provided with AIC, AICC and BIC statistics in Tables 36 and 37 of the CS (B.3.3). These show that for both study arms, the lognormal distribution provides the best fit to PFS data, followed by generalised gamma and loglogistic distributions.
- **Plausibility of extrapolation:** The company states that plausibility was assessed by visual inspection of the curves by oncologists currently practising in the NHS and England based on their clinical experience. No further information is provided about how this clinical assessment of validity was conducted or how it informed the choice of PFS curve.

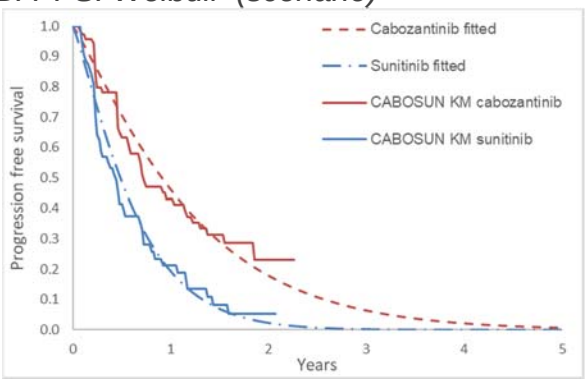
A: PFS: Exponential (scenario)



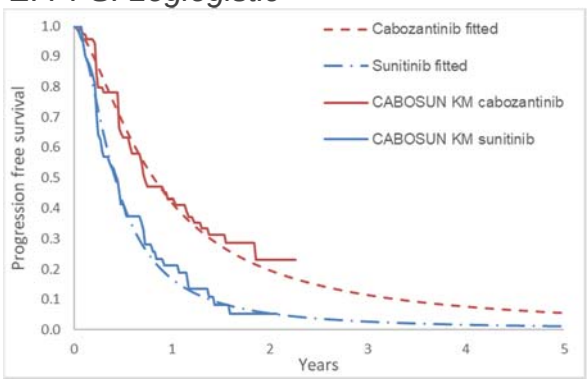
D: PFS: Lognormal ***(base case)***



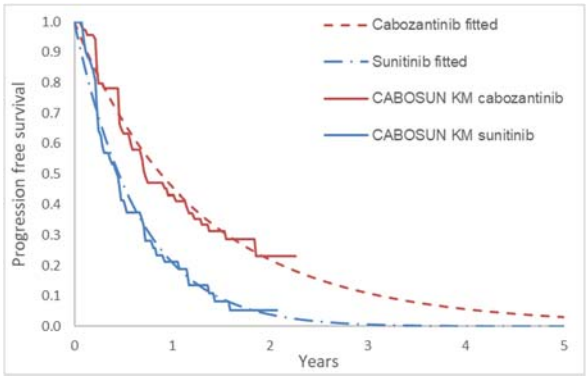
B: PFS: Weibull (scenario)



E: PFS: Loglogistic



C: PFS: Gompertz (scenario)



F: PFS: Generalised gamma

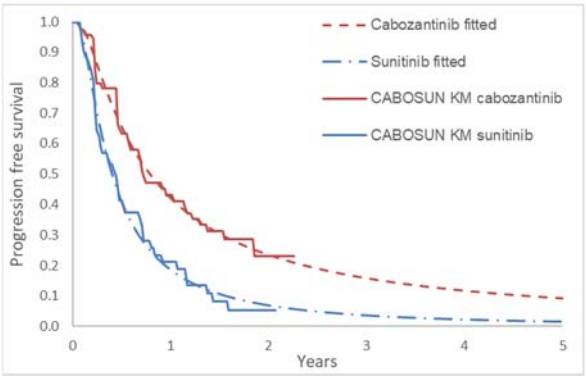


Figure 17 PFS curves - fitted to CABOSUN data (direct comparison)

Source: Figures generated by ERG from company model and KM data

We agree that the lognormal, exponential and Gompertz functions show a reasonable visual fit to the trial data, although they both overestimate median PFS for cabozantinib (as do all the other functions due to a 'dip' in the PFS KM curve). The Weibull has a poor visual fit.

PFS indirect comparison

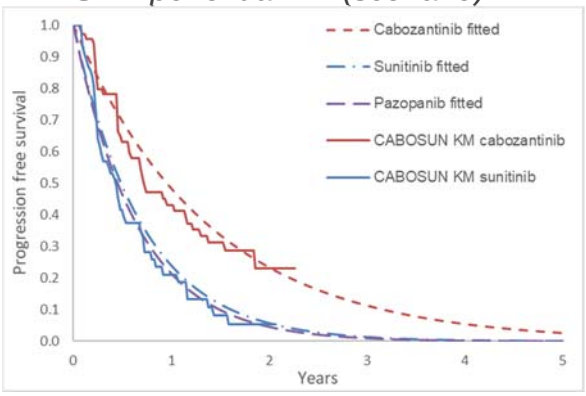
ITC curves for PFS are shown in Figure 18 alongside the CABOSUN KM plots.

The company use the second-order FP model with P1=P2=-1 for their ITC base case, and exponential, Weibull and Gompertz random effect models in scenario analyses. This choice was based on the following considerations:

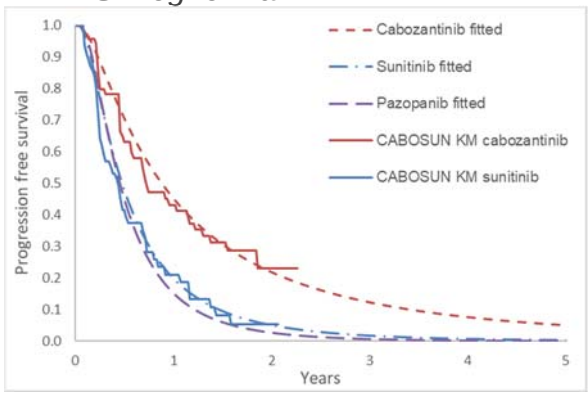
- **Proportional hazards (PH):** The company concludes that the PH assumption was violated for PFS in both CABOSUN and COMPARZ (Appendix D, Table 12). As noted above, we question this conclusion for CABOSUN. But for COMPARZ, the Schoenfeld and log cumulative hazard plots do suggest a change in hazard ratio over time (section 3.1.7.3 above).
- **Goodness-of-fit (DIC):** Evidence of the fit of the parametric and FP curves to the CABOSUN and COMPARZ data is provided with DIC statistics in Tables 23 and 24 of the CS (section B.2.9). For the parametric models, the lognormal distribution had the lowest DIC with both, followed by loglogistic. Among the FP models, the second-order P1=P2=-1 model had the lowest DIC statistic.
- **Plausibility of extrapolation:** The company notes that they decided not to include other second-order FP models in scenario analyses because they predict a high PFS rate at year 5. There is no discussion of the plausibility of the other parametric models, or of the long-term continuation of treatment effects observed in the trials.

Summary statistics in Table 14 show that all the ITC models overestimate median PFS for cabozantinib in relation to the CABOSUN result and several also overestimate median PFS with sunitinib. Long-term projections of PFS also seem optimistic for some models, particularly the second-order FP models and the Gompertz parametric model. Although the exponential and Weibull models do not have this problem, they have a worse visual fit and large overestimates of median PFS. We conclude that the lognormal and loglogistic models seem to provide the best balance of fit to the CABOSUN data with a realistic long-term extrapolation.

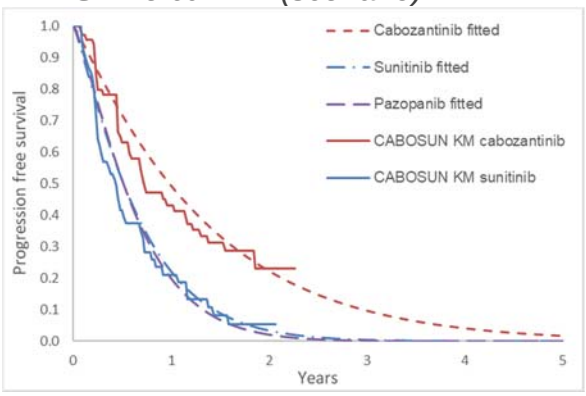
A: PFS: Exponential RE (scenario)



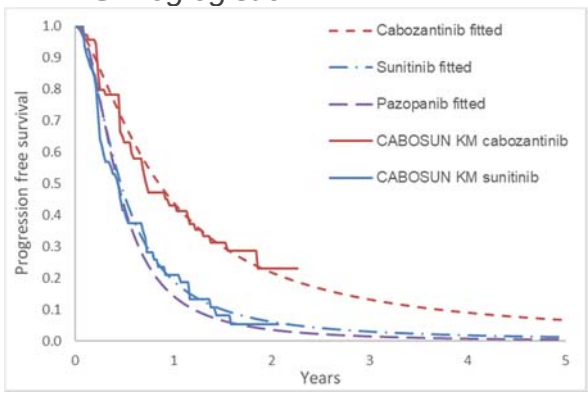
D: PFS: Lognormal RE



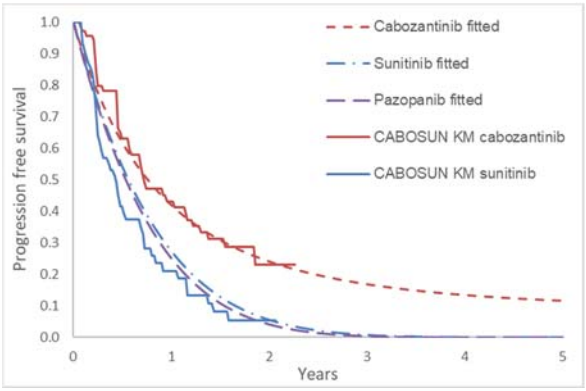
B: PFS: Weibull RE (scenario)



E: PFS: Loglogistic RE



C: PFS: Gompertz RE (scenario)



F: PFS: FP P1=P2=-1 ***(base case)***

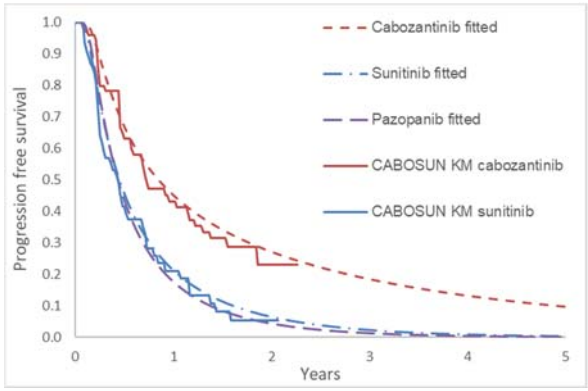


Figure 18 PFS curves – ITC models fitted to CABOSUN AND COMPARZ

Source: Figures generated by ERG from company model and KM data

Table 14 Summary statistics for PFS curves

	Model fit statistics [a]		Median PFS (months)			5-year PFS (%)		
			Cabo	Suni	Pazo	Cabo	Suni	Pazo
Data sources								
CABOSUN	-		8.6	5.3	-			
COMPARZ [b]	-		-	9.5	8.4			
IMDC data [c]	-		-	8.4	8.3	-	5%	8%
Tivozanib STA[d]	-		-	6.8	8.4			
Fitted models: direct comparison (CABOSUN)								
Exponential	325.6	303.2	10.9	5.6	-	2%	0%	-
Weibull	328.5	304.8	11.2	6.1	-	1%	0%	-
Gompertz	329.8	307.3	10.9	5.7	-	3%	0%	-
Lognormal	321.7	295.8	10.2	5.4	-	5%	0%	-
Loglogistic	323.6	298.0	10.0	5.3	-	5%	1%	-
Gamma	324.6	298.8	9.7	5.1	-	9%	1%	-
Fitted models: ITC parametric random effects (CABOSUN & COMPARZ)								
RE Exponential	1941.6		11.7	6.1	5.6	3%	0%	0%
RE Weibull	1945.5		12.0	6.4	6.3	2%	0%	0%
RE Gompertz	1943.5		9.3	7.0	6.6	11%	0%	0%
RE Lognormal	1860.6		10.8	5.9	5.6	5%	0%	0%
RE Loglogistic	1887.8		10.4	5.7	5.4	7%	1%	1%
Fitted models: ITC fractional polynomials (CABOSUN & COMPARZ)								
FP P=-1	1910.4		11.9	6.6	6.4	1%	0%	0%
FP P=-0.5	1932.0		12.0	6.7	6.5	1%	0%	0%
FP P=0	1945.9		11.7	6.6	6.5	2%	0%	0%
FP P=0.5	1947.6		11.4	6.4	6.3	3%	0%	0%
FP P=1	1943.6		11.2	6.2	5.9	6%	0%	0%
FP P1=-0.5, P2=0	1852.1		10.5	5.4	5.3	14%	4%	3%
FP P1=-1, P2=0	1840.3		10.5	5.3	5.2	12%	3%	2%
FP P1=-1, P2=-1	1825.0		10.4	5.6	5.4	10%	0%	0%
FP P1=-1, P2=0.5	1850.4		10.3	5.5	5.3	17%	4%	3%
FP P1=-1, P2=1	1858.1		10.2	5.7	5.5	21%	5%	3%

a As reported in CS Tables 23, 24, 34 and 35: direct comparison Bayesian Information Criterion (BIC) for cabozantinib / sunitinib; parametric ITC: Deviance Information Criterion (DIC) for fixed/random effects models; and ITC FPs: Deviance Information Criterion (DIC) for first/second order models

b Motzer et al. 2013 analysis of COMPARZ trial data.[34]

c Ruiz-Morales et al. analysis of 7438 patients with metastatic RCC treated at first line with sunitinib (n=6519) or pazopanib (n=919).[47]

d ERG preferred results from Tivozanib STA (TA512)[22]

4.3.4.3 Time to discontinuation (TTD)

Time on treatment is modelled for cabozantinib and sunitinib based on parametric curves fitted to CABOSUN data (CS B.3.3 Figure 15). For pazopanib, no TTD data were available, so the company assume that the sunitinib curve would also apply to pazopanib. They justify this by noting that the mean treatment duration in COMPARZ was 11.5 months for both treatments. The median duration of treatment was also similar: 7.6 months (range 0 to 38) for sunitinib and 8.0 months for pazopanib (range 0 to 40) respectively.[35] The ERG agrees with this approach.

We illustrate the fitted TTD curves alongside KM plots (digitised by the ERG) in Figure 19 below. The visual fit appears similar for the different parametric functions and as the TTD data are mature, extrapolation is less of an issue than for PFS and OS. The model fit statistics are shown in Table 38 and 39 of the CS (B.3.3). The optimum curve differs by measure of fit and by treatment. The company use the lognormal in their base case analysis, as they argue that this provides a good fit for both cabozantinib and sunitinib. They also test exponential, Weibull, Gompertz and generalised gamma in scenario analysis. There is no obvious reason for excluding the loglogistic from scenario analysis.

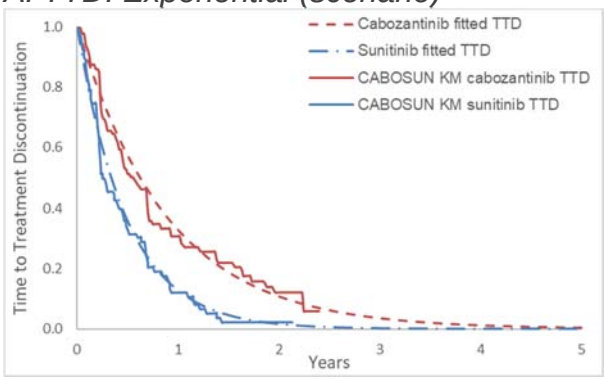
Table 15 below summarises the survival curves used for OS, PFS and TTD.

Table 15 Survival curves used in company analyses

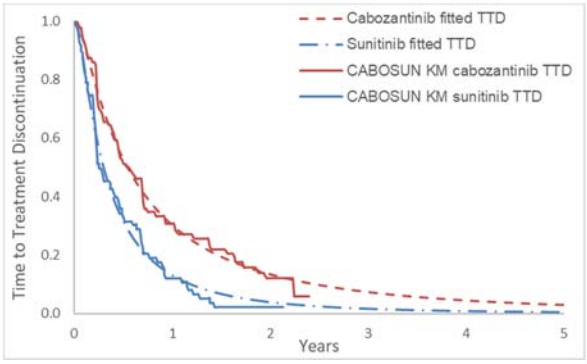
Curve	Method	Treatment	CS Base case	CS scenarios
OS	**Direct** CABOSUN	Cabozantinib Sunitinib	Exponential	Weibull Gompertz
	ITC CABOSUN & COMPARZ	Cabozantinib Sunitinib Pazopanib	FP P1=P2=-1	RE exponential RE Weibull RE Gompertz FP P1=-0.5, P2=0 FP P1=-1, P2=0
PFS	**Direct** CABOSUN	Cabozantinib Sunitinib	Lognormal	Exponential Weibull Gompertz
	ITC CABOSUN & COMPARZ	Cabozantinib Sunitinib Pazopanib	FP P1=P2=-1	RE exponential RE Weibull RE Gompertz
TTD	**Direct** CABOSUN	Cabozantinib Sunitinib Pazopanib	Lognormal	Exponential Weibull Gompertz Generalised gamma

RE = Random effects; FP = Fractional polynomial

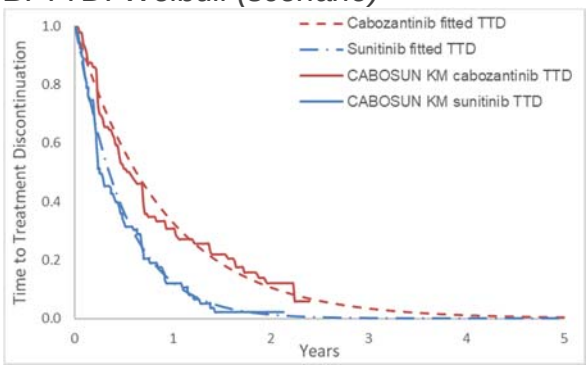
A: TTD: Exponential (scenario)



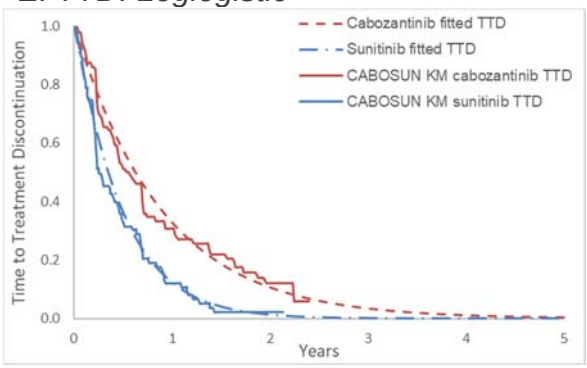
D: TTD: Lognormal ***(base case)***



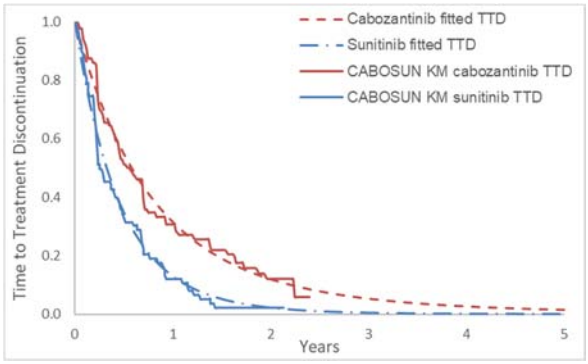
B: TTD: Weibull (scenario)



E: TTD: Loglogistic



C: TTD: Gompertz (scenario)



F: TTD: Generalised gamma (scenario)

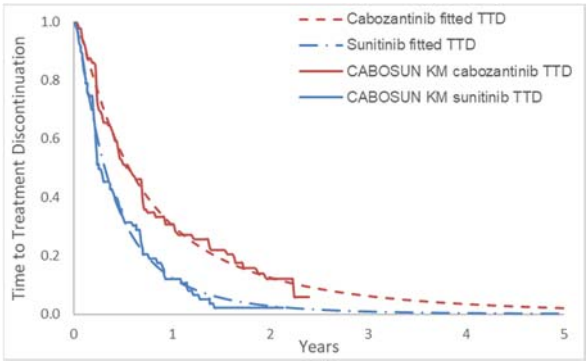


Figure 19 TTD curves - fitted to CABOSUN data (direct comparison)

Source: Figures generated by ERG from company model and digitised KM from CS B.3.3 Figure 15

4.3.5 Health related quality of life

For calculation of QALYs, the model requires estimates of utilities for the two health states (PF and PD) and disutilities for grade 3 or 4 treatment emergent adverse events (TEAEs).

EQ-5D or other relevant utility data was not collected in CABOSUN. The CS describes a post hoc analysis using a quality-adjusted time without symptoms of disease or toxicity (Q-TWiST) framework. Outputs from this analysis are not fed into the economic model. As they fall outside the NICE reference case, we do not discuss them further.

The company conducted a systematic search to identify utility values that could be used in the model. The company's search strategy is described in CS Appendix H. The ERG considers this search strategy to be adequate and up to date (see section 3.1.1). The search identified 22 publications which reported EQ-5D-based utilities relevant to first-line treatment of advanced or metastatic RCC. Of these publications, the company deemed four to be relevant to a UK setting; summarised in CS Table 40. Of these four publications, only Swinburn et al. 2010[15] reported health state-specific utilities that would be suitable for inclusion in the model. EQ-5D utility values reported in Swinburn et al. 2010 include; stable disease with no adverse event (0.795) and disease progression (0.355). In addition, the company checked and reported relevant utility values from previous NICE submissions. The utility values used in the company base case and scenarios are summarised in Table 16 below.

Table 16 Utility values (adapted from CS Tables 46 & 47)

Health state	Utility value: mean (SE)	95% CI	CS reference
Progression free			
Base case	0.726 (0.011)	0.705 to 0.748	Tivozanib TA512 [22]
Scenario	0.70 (0.01)	0.680; 0.720	Pazopanib TA215[48]
Scenario	0.78 (0.078)*	0.627; 0.933*	Sunitinib TA169[49]
Scenario	0.795 (0.0176)	0.761; 0.830	Swinburn 2010[15]
Progressed disease			
Base case	0.649 (0.019)	0.612 to 0.686	Tivozanib TA512 [22]
Scenario	0.59 (0.059)*	0.474; 0.706*	Pazopanib TA215[48]
Scenario	0.705 (0.071)*	0.567; 0.843*	Sunitinib TA169[49]
Scenario	0.355 (0.0288)	0.299; 0.412	Swinburn 2010[15]
TEAE grade 3/4			
Base case	-0.2044 (0.0682)	-0.0707 to -0.3381	COMPARZ[50]
Scenario	-0.0550 (0.0068)	-0.0418; -0.0685	METEOR trial[51]

Abbreviations: TEAE, treatment emergent adverse effects; CI, confidence interval; SE, standard error; *SE or 95% CI not available in literature; 10% of the mean assumed.

4.3.5.1 Health state utilities

The ERG agrees that the health state utility values in Table 16 meet the NICE reference case and are suitable for inclusion in the model. The values for the progression free health state are reasonably consistent (0.70 to 0.80). However, there are large differences between the available utility estimates after disease progression (0.36 to 0.71). The Swinburn et al.[15] study gives the biggest difference in utility between the PF and PD state, a loss of 0.44. This study used time trade-off approach to elicit UK societal preferences for health states associated with metastatic RCC, rather than EQ-5D valuations. The ERG agrees with the company's preference for the health state utility values used in the tivozanib STA, with scenarios based on the pazopanib and sunitinib STA utility values and Swinburn et al.[15]

We spotted a disparity between the utility scores from the sunitinib STA as reported in CS Tables 43 and 47 and the values applied in the company's model. This error did not affect the results for the company's base case, but the results for the scenario analysis with sunitinib health state utilities in CS Tables 66 and 67 are incorrect. The company corrected these errors in response to a clarification question (B4).

4.3.5.2 Adverse event disutilities

As adverse event specific utility data was not reported in CABOSUN, the company derived base case TEAE disutilities from the COMPARZ study, with values from the METEOR trial[51] in scenario analyses. The company assumes that TEAE disutilities are not disease-specific and that all types of grade 3 or 4 events elicit the same utility loss for a fixed period of 4 weeks and a fixed number of episodes per patient per TEAE. These assumptions yield a mean QALY loss of 0.0225 per TEAE in the base case (0.006 in the METEOR trial-based scenario). 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 analyses to test model sensitivity to the TEAE disutility parameter, including higher as well as lower estimates of the disutility.

The company models the incidence of grade 3/4 TEAEs based on reported rates from the CABOSUN trial for cabozantinib and sunitinib and from COMPARZ for pazopanib, see Table 17 below. Only events with a reported incidence of 5% or greater in at least one arm were included. 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 QALY decrements for TEAEs associated with second line treatments. We consider that this is reasonable, as utility loss related to subsequent treatments should be reflected in the PD health state utility. This may also be true for the PF health state utility – thus there may be some degree of double counting due to the inclusion of TEAE disutilities for the first-line treatments. However, it is important to reflect potential disutility related to differences in adverse effect incidence for main treatments of interest.

Table 17 Incidence of modelled grade 3/4 adverse events by treatment and study

Adverse Event	Cabozantinib (n=78) CABOSUN (%)	Sunitinib (n=72) CABOSUN (%)	Pazopanib (n=554) COMPARZ (%)
Decreased appetite	5	1	1
Diarrhoea	10	11	9
Dyspnoea	1	6	3
Embolism	8	-	-
Fatigue	6	17	11
Hyperglycaemia	-	6	5
Hypertension	28	21	15
Hyponatremia	9	8	7
Hypophosphatemia	9	7	4
Hypotension	5	1	-
Increased ALT	5		17
Increased AST	3	3	12
Lymphocytopenia	1	6	5
Neutropenia	-	3	5
Pain	5	-	-
Hand and foot syndrome	8	4	6
Stomatitis	5	6	1
Thrombocytopenia	1	11	4

Source: Extracted from the model by the ERG.

4.3.6 Resource use and costs

The costs included in the economic model consist of drug acquisition and administration costs for first-line and subsequent treatments (adjusted for dose intensity and wastage where appropriate), health state management costs (for PF and PD), costs incurred for the management of adverse event costs and costs incurred at the end of life.

The company conducted a systematic literature review to identify published resource use and cost data relevant to the cost-effectiveness analysis. From a total of 61 full-text articles identified, the company judged 22 studies to be eligible for data extraction since they related to countries in the company's scope (European countries, Australia and Canada). CS Appendix I provides a detailed description of the company's search strategy and inclusion criteria. The ERG considers that the company's literature review, which was updated in September 2017, is likely to reflect available evidence.

4.3.6.1 First-line drug costs

Table 18 summarises the drug acquisition costs for first-line treatments included in the company's model.

Table 18 Drug cost per week for first line treatments (adapted from CS Table 48)

Drug	Relative dose intensity (SE)	PAS discount	Cost per week without discount	Cost per week with discount
Cabozantinib	94.3% (1.5) [a]	[redacted]	[redacted]	[redacted]
Sunitinib	87.4% (6.3) [a]	First 6-week cycle free [c]	£457	£457
Pazopanib	86.0% (8.6) [b]	12.5% [b]	£450	£394

a CABOSUN CSR, Table 37, sunitinib mean relative dose intensity.[27]
b NICE pazopanib appraisal TA215[13]
c NICE sunitinib appraisal TA169.[12] Coded in model rather than as a simple discount

The cost per pack for all drugs are derived from the British National Formulary. The company base case includes published patient access schemes (PAS) discounts for pazopanib and sunitinib: 12.5% reduction on the list price for pazopanib (TA215)[13] and the first 6-week cycle free for sunitinib (TA169).[12] The company also applies a pre-existing confidential PAS discount of [redacted] for cabozantinib in previously-treated advanced RCC (TA463[14]) to the cost of first-line cabozantinib in their base case analysis.

Relative dose intensity is factored into the cost calculations to reflect the percentage of days with interrupted treatment – for example, due to adverse effects. Cabozantinib is available in doses of 20, 40 and 60 mg, with all doses priced equally. However, while a reduction in dose

does not impact on costs, an interruption in treatment may do so. The company bases estimates of dose intensity for cabozantinib and sunitinib on CABOSUN data and for pazopanib, estimates in the NICE appraisal TA215 are used. For comparison, the ERG analysis for the recent NICE tivozanib appraisal (TA512) included an 86% relative dose intensity for both sunitinib and pazopanib, based on values cited in previous NICE appraisals for these drugs. The appraisal committee in TA512 concluded that there is uncertainty over the impact of dose intensity on the cost of oral treatments, and that this is likely to fall somewhere between the ERG's estimates and 100%. We consider that the company's approach to modelling the cost impact of dose intensity is reasonable. However, we conduct an additional scenario analysis to test the effect of assuming the same relative dose intensities (86% and 100%) for all treatments.

The company did not include additional administration costs for oral chemotherapies in their model. The ERG agrees with this approach. We note that the NHS does incur costs for delivery of oral chemotherapies, which are included the National Schedule of Reference Costs (currency code SB11Z). However, the modelled health state costs (listed below) include a monthly consultant-led medical oncology outpatient visit and blood tests, which we assume would include the cost of procurement, prescribing and monitoring of oral chemotherapies.

4.3.6.2 Health-state costs

The CS reports assumptions about resource use and unit costs for disease management in Tables 49 and 50 (summarised in Table 19 below). The company assumes that patients have a monthly medical oncologist visit and blood tests and a computerised tomography (CT) scan every three months. For scenario analysis, they assume less frequent oncologist follow-up but with access to a specialist nurse. The company's model makes provision for second-line treatment following treatment failure (see below), with the same follow up and monitoring pre and post-progression.

The ERG considers that the company's estimates of health state costs are reasonable. They reflect resource use assumptions in previous NICE appraisals[12 13 22] and experts consulted by the ERG did not object to the company's assumptions, except that it was noted that in routine NHS care, patients would have some follow-up with a nurse specialist. Unit costs are based on appropriate and up-to-date national sources.[52 53]

Table 19 Health state management costs (adapted from CS Table 49 and 50)

Health state	Resource	Frequency per week		Unit cost	Reference
		Base case	Scenario		
PF	Outpatient (first)	Not applicable		£219	NHS Reference Costs 2016/17. Currency code WF01B, Service code 370, Medical oncology
	Outpatient (follow up)	0.25	0.08	£173	NHS Reference Costs 2016/20. Currency code: WF01A, Service code 370, Medical oncology
	Nurse visit	0	0.25	£173	Cost per hour. Nurse (GP practice), PSSRU Unit costs of health and social care 2016
	CT scan	0.08	0.08	£115	NHS Reference Costs 2016/17. Currency code: RD25Z. CT of three areas, without contrast
	Blood test	0.25	0.25	£3	NHS Reference Costs 2016/17. Currency code: DAPS05
PD	Outpatient (follow up)	0.25	0.08	£173	NHS Reference Costs 2016/17. Currency code: WF01A, Service code 370, Medical oncology
	Nurse visit	0	0.25	£173	Cost per hour. Nurse (GP practice), PSSRU Unit costs of health and social care 2016
	CT scan	0.08	0.08	£115	NHS Reference Costs 2016/17 Currency code: RD25Z Computerised Tomography Scan of three areas, without contrast
	Blood test	0.25	0.25	£3	NHS Reference Costs 2016/17. Currency code: DAPS05

Abbreviations: PF, progression free; PD, progressed disease; CT, computerised tomography.

4.3.6.3 Adverse event costs

The model includes costs for managing grade 3/4 TEAE with an incidence of ≥5% in the CABOSUN (cabozantinib and sunitinib) and COMPARZ (pazopanib) trials (see Table 17 above). The company's base case estimates of the costs of managing these events are summarised in Table 20. Resource use assumptions were derived from published estimates (CS Appendix I), HTA reports and clinical opinion. To avoid double counting, the company model omits costs for adverse events such as hyponatremia and hypotension which would be managed as part of regular follow up (included in the health state costs listed in the previous section). Unit costs come from standard national sources: including NHS Reference costs 2016/17, the British National Formulary (8/12/17) and the PSSRU Unit costs of Health and Social Care report 2016 (see Table 21).

The ERG finds the resources included in the CS to be comprehensive. We spotted certain textual errors in the CS: the unit cost for vascular ultrasound scan is wrongly reported as £75 and the cost of hospitalisation for lymphocytopenia is reported as £429. However, these errors do not affect cost-effectiveness results since the correct values are used in the model.

Table 20 Costs for management of adverse events (Adapted from CS Table 53)

Adverse event	Cost per event (£)	Assumptions about resource use
Diarrhoea	£567	Based on pazopanib NICE STA. Short stay admission and Loperamide 2 mg102 q.i.d 30 days
Dyspnoea	£68	Based on assumption of one pulmonologist visit
Embolism	£1,640	Based on NICE guidance on venous thromboembolic diseases: 1 ultrasound of coronary vessels. Therapy initiation with low molecular weight heparin for 6 months: deltaparin 18000 units o.d. units for first 30 days and continue with deltaparin 15000 units o.d. for further 5 months
Fatigue	£35	Based on tivozanib NICE STA. 20% of patients will have additional outpatient attendance
Hyperglycaemia	£156	Based on assumption: 1 visit to endocrinologists. Initiation of therapy with p.o anti-diabetic medication: metformin 500mg3 o.d. for one year
Hypertension	£128	Based on tivozanib NICE STA. 3 GP attendances, ramipril 5 mg + bendroflumethiazide 2.5 mg o.d. for 1 year
Lymphocytopenia	£362	Based on assumption of 20% of short stay emergency tariff (weighted average of SA35A-SA35E: £ 515) and 80% of day case tariff (weighted average of SA35B-SA35E: £ 288)
Neutropenia	£1,107	Based on the assumption: Granulocyte colony-stimulating factors (granulocyte CSF): Filgrastim. 5µg/kg for 14 days (dose is 450 µg o.d. for TM=90kg) Neupogen 30million units/1ml (1µg=100000 units)
Pain	£138	Based on assumption: Outpatient visit for pain management (CS Table 53 incorrectly cites monthly visits, but only one is costed in model).
Hand and foot syndrome	£104	Based on tivozanib NICE STA: 60% of patients will have additional outpatient attendance
Stomatitis	£42	Based on assumption: Local therapy for pain relief, local anaesthetics or other anti-inflammatory preparations - oral solution of dexamethasone 2mg/5ml
Thrombocytopenia	£351	Based on assumption: 20% of short stay emergency tariff (weighted average of SA12G-SA12K) and 80% of patients with day case tariff (weighted average of SA12G-SA12K)
Hyponatremia	£0	

Adverse event	Cost per event (£)	Assumptions about resource use
Hypophosphatemia	£0	Regular blood tests already considered under disease Management costs
Increased ALT	£0	
Increased AST	£0	
Hypotension	£0	Monthly outpatients visit already covered by disease management costs
Decreased appetite	£0	No stated justification in CS but not associated cost in company's model

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; od, once daily; qid, four times a day; STA, single technology appraisal

Table 21 Unit costs for management of adverse events (CS Table 51)

Health resource	Cost, £	Reference
Short stay admission due to diarrhoea	£558	HRG FD02E, Inflammatory Bowel Disease without Interventions, with CC Score 5+, NHS reference costs 2016/2017 [53]
Vascular ultrasound scan	£65 *	HRG RD47Z, NHS reference costs 2016/2017
Outpatient attendance for hand and foot syndrome	£173	HRG WF01A: service code 370 Medical oncology, NHS reference costs 2016/2017
Visit to endocrinologist due to hyperglycaemia	£146	WF01A, Service code 302, Endocrinology, NHS reference costs 2016/2017
GP visit due to hypertension	£36	GP visit-Unit cost per surgery consultation; PSSRU Cost of health and social care 2016[53]
Hospitalisation cost due to lymphocytopenia - Short stay emergency tariff	£492 *	HRG SA35A-SA35E short stay emergency tariff (weighted average), NHS reference costs 2016/2017
Hospitalisation cost due to lymphocytopenia - Day case	£330	HRG SA35A-SA35E day case tariff (weighted average), NHS reference costs 206/2017
Hospitalisation cost due to thrombocytopenia - Short stay emergency	£522	HRG SA12G-SA12K short stay emergency tariff (weighted average), NHS reference costs 2016/2017
Hospitalisation cost due to Thrombocytopenia - Day case	£308	HRG SA12G-SA12K day case tariff (weighted average), NHS reference costs 2016/2017
Outpatient attendance due to dyspnoea	£68	WF01A, Service code 342, Programmed Pulmonary Rehabilitation, NHS reference costs 2016/2017
Outpatient visit for pain management	£138	WF01A, Service code 191, Pain management, NHS reference costs 2016/2017

* Values from CS Table 51 corrected by ERG. Correct values were used in the submitted company model.

4.3.6.4 Second line treatment use and costs

The company's assumptions about the proportions of patients receiving subsequent treatments after failure of initial therapy are reported in Table 22. The base case reflects second-line treatments in the trials that inform model survival parameters: CABOSUN for cabozantinib and sunitinib; and COMPARZ for pazopanib. The company notes that although some of these treatments are not available or not approved for second line use in England, costing this mix of subsequent treatments is consistent with the implicit inclusion of their benefits through the trial estimates of survival. The company also conduct a scenario analysis in which the cost of subsequent treatment is adjusted to better reflect NHS practice. This scenario is largely based on ERG assumptions that were made for the NICE tivozanib appraisal.[22] For this current appraisal, the company adjust the distribution after first-line treatment with sunitinib or pazopanib, assuming that 10% of patients would start cabozantinib but only 40% axitinib.

Table 22 Distribution of subsequent treatments (Adapted from CS Table 56 and 57)

Second-line treatments	Following initial treatment with:					
	Company base case			Scenario analysis		
	Cabo[a]	Suni[a]	Pazo[b]	Cabo[c]	Suni[c]	Pazo[c]
Axitinib	23%	19%	6%	50%	40%	40%
Pazopanib	16%	12%	0%	0%	0%	0%
Sunitinib	13%	13%	29%	0%	0%	0%
Temsirolimus	9%	4%	6%	0%	0%	0%
Nivolumab	13%	15%	0%	30%	30%	30%
Everolimus	8%	19%	31%	10%	10%	10%
Sorafenib	1%	3%	11%	0%	0%	0%
Bevacizumab	0%	6%	7%	0%	0%	0%
Cabozantinib	1%	6%	0%	0%	10%	10%
Interferon	1%	0%	0%	0%	0%	0%
BSC	0%	0%	0%	10%	10%	10%

Abbreviations: cabo, cabozantinib; suni, sunitinib; pazo, pazopanib; BSC, best supportive care Sources:

a CABOSUN Clinical Study Report, Table 26.[27]
b COMPARZ Clinical Study Report [38]
c Tivozanib NICE STA[22] with 10% utilisation moved from axitinib to cabozantinib (following first-line treatment with sunitinib or pazopanib).

The ERG agrees with the company's general approach to modelling second-line treatments, with the base case following utilisation in the clinical trials and scenario analysis testing costs that are more reflective of NHS practice. This is not ideal, as the scenario omits the impact of NHS practice on survival, and the direction and magnitude of bias from this omission is unclear. However, we do not believe that it is feasible to model the effects of a different mix of second-line treatments. Re-analysis of OS data to adjust for second line treatment would not

be possible given the small sample size in CABOSUN: for example, only 4 patients in the cabozantinib arm and 6 in the sunitinib arm received nivolumab as subsequent therapy.[27] An alternative would be to explicitly model survival for the different second-line treatments after discontinuation of the initial therapy, but this would require a new model and systematic evidence review for the relevant population.

The mix of second-line treatments in the company's scenario analysis is similar to that used in the recent NICE appraisal for tivozanib. It excludes treatments that are not recommended by NICE for second-line use (sunitinib and sorafenib)[20] and those that have not been appraised for this indication (pazopanib, temsirolimus, bevacizumab and interferon-α). It includes three treatments recommended by NICE for second-line use: cabozantinib (TA463), nivolumab (TA417) and everolimus (TA432).[54] However, it does not include lenvatinib plus everolimus, which was also recently approved by NICE for second line use (TA498).[55] Clinical advice to the ERG is that current practice is usually to offer pazopanib or sunitinib at first-line, followed by nivolumab or cabozantinib at second line. If cabozantinib is made available at first-line, nivolumab or lenvatinib and everolimus would then probably be offered at second line. We conduct additional scenario analyses to test the impact of changing the costs of second line treatments.

Table 23 below (adapted from CS Tables 55 and 58) summarises the costs for second-line treatments included in the model. The company costs all second-line treatments at list price. This does not reflect current prices paid by the NHS, because agreed PAS discounts are in place for the five treatments approved by NICE for previously-treated advanced renal cell carcinoma. In ERG additional analyses below, we use the same approach and do not include PAS discounts for any second-line treatments. We apply all available PAS discounts in a confidential addendum to this report.

As for first-line treatments, the company assume that oral treatments at second-line do not incur an additional administration cost. We consider that this is reasonable, given that health management costs continue to include monthly outpatient visits after disease progression. The model includes administration costs for drugs delivered by injection (interferon-α) and IV infusion (nivolumab, bevacizumab and temsirolimus). The company assume that 25% of interferon-α injections are administered by a district nurse at a cost of £37 (£9.25 per dose). CS Table 59 states that the cost per IV infusion is £199, but the model actually applies a cost of £205 per infusion for bevacizumab and temsirolimus and omits the administration cost for nivolumab. We believe that this is an error and apply a cost of £205 per infusion for all drugs

delivered by IV infusion: this is the 2016/17 NHS Reference Cost for outpatient delivery of subsequent elements of the chemotherapy cycle (currency code SB15Z).

Table 23 Costs and duration of subsequent treatments (Adapted from CS Tables 55 and 58)

Subsequent treatments	Relative dose intensity, % (SE)	Cost per week, list price [a]	Duration (SE), weeks
Axitinib	102 (1.9)	£897	31.5 (3.2)
Pazopanib	86 (8.6)	£450	49.8 (5.0)
Sunitinib	87 (6.3)	£457	24.7 (2.5)
Temsirolimus	92 (9.2)	£825	17.0 (1.7)
Nivolumab	98 (9.8) plus 8% wastage	£1,572 *	42.0 (4.2)
Everolimus	84 (1.1)	£523	23.9 (2.4)
Sorafenib	80 (8.0)	£715	25.8 (2.6)
Cabozantinib	93.3 (9.3)	[redacted]	33.1 (3.3)
Bevacizumab	88 (8.8)	£1,050	24.0 (2.4)
Interferon-α	86 (8.6)	£155	12.0 (1.2)

[a] Cost at list price, including: adjustment for dose intensity; administration cost (£205 per infusion; £ and drug wastage (for drugs delivered by IV infusion or injection).

* Excludes cost of administration. This is corrected in ERG analysis.

The costs in Table 23 are also adjusted for relative dose intensity to account for missed doses of medications and wastage for vial formulations (nivolumab, bevacizumab and temsirolimus). The company also conducts a scenario analysis without wastage. We consider that wastage is likely to occur in clinical practice with vial formulations, and so should be included in the analysis.

The duration of subsequent treatments is based on a variety of sources, including NICE STAs for axitinib (TA333)[56] and nivolumab (TA417),[57] CABOSUN and COMPARZ for sunitinib and pazopanib, METEOR for cabozantinib and other trials for temsirolimus, everolimus, bevacizumab.[58 59 60]

4.3.6.5 End of life costs

The company base case includes a cost for end of life care applied in the last cycle before death. This comes from a 2014 Nuffield Trust report that estimated the cost of hospital care in the last three months of life for people within 2 years of a cancer diagnosis at £5,890. The company uprated this to £6208 at 2017 prices, based on general inflation indices.[61 62] The ERG believes that this is an under-estimate, due to the omission of costs for local-authority funded social care, district nursing and GP visits and the company's method of adjusting for inflation. Based on the Nuffield report, we estimate an end of life cost of £7,961 from an NHS

and PSS perspective and inflating using the Hospital and Community Health Services price index.[52] We include this revised figure in ERG analyses, but also conduct a scenario analysis excluding end of life care costs.

4.3.7 Model validation

The company state that model outputs were validated by UK clinical oncologists (CS B.3.10). No details are given about how this validation process was done or whether any changes were made as a result. It is also stated that the model was verified by economists not involved in its development. A list of verification checks is given, including checks on input data and technical validation of coding.

4.3.7.1 ERG model verification procedures

We conducted a range of manual checks to verify model inputs and calculations ('white box' tests) and to test the face-validity of the model results ('black box' checks):

- Cross-checking of all parameter inputs against values in the CS and cited sources;
- We traced input parameters from entry cells in the model ('User-Inputs" and "Resources" sheets), to PSA/DSA sampling (on the "Variables" sheet) through to the survival curve and Markov calculation sheets;
- We independently replicated calculations for first and second line drug costs (to check adjustments for dose, intensity, wastage and PAS discounts), health state costs and adverse event costs and QALY loss;
- Survival curve calculations were checked ("TPs_CABOSON", "TPs_ITC" and "TPs_ITC_FP" sheets).
- Use of PFS, OS and TTD results to estimate the distribution of the cohort by health state and the numbers of events over time in Markov trace sheets (E.Cabo.RCT etc.)
- We checked QALY and cost calculations on the Markov sheets;
- And the links from the total costs and outcomes on the Markov sheets back to the ICER calculations on the 'Results' sheet.
- We checked all model outputs against results cited in the CS, including the base case, PSA and DSA and we manually ran scenarios.

Through this process we identified some errors and inconsistencies:

1. QALY calculations – discounting and utility adjustment for age and sex were applied twice in the Markov trace sheets. This had the effect of shrinking the estimated QALYs for all treatments and hence the incremental QALY differences between treatments, thus overestimating the correct ICER.

2. QALY calculations – the one-off QALY loss that is applied in the first cycle to account for adverse effects of treatment was incorrectly adjusted for the duration of the cycle. The reduced the effect of adverse events on QALYs for all treatments.
3. The 'Accrual Utility' column in the Markov sheets also adjusted the QALYs accrued in each cycle again for the duration of the cycle. Thus the graphs of cumulative QALYs over time on the 'Table1' sheet are incorrect. This does not influence the cost-effectiveness results.
4. Utility estimates for PF and PD health states from the sunitinib NICE technology appraisal that were used in scenario analyses were incorrectly entered in the model. This was corrected by the company in response to a clarification question.
5. The cost of administering the nivolumab infusion was not included in calculation of second-line treatment costs. The unit cost for this administration cited in the model was also different for nivolumab than for bevacizumab and temsirolimus, which are also administered by infusion.
6. We believe that the cost of end of life care was incorrectly estimated for the NHS and PSS perspective and that it was incorrectly updated for inflation.
7. There was an error in the scenario analysis 'PFS=OS=Gompertz' (CS Table 66) (the PFS curve was set to lognormal rather than Gompertz).
8. The 'health resource (UK clinicians)' scenario analysis (CS Tables 66 and 67) gave the base case ICER because of an error in the linking of the Source of health resource control on the User_Input sheet.
9. On the 'Curve data' sheet, TTD for pazopanib was calculated as a proportion of pazopanib PFS, rather than being set equal to sunitinib TTD. This did not influence the cost-effectiveness results, although the summary statistics and graph for TTD on the 'User-inputs' sheet are incorrect.

4.3.7.2 Assessment of internal and external validity of model

Key statistics relating to the fit of the company's fitted survival models are shown in Table 13 and Table 14 for OS and PFS respectively. In addition to the model fit statistics (BIC and DIC), we show median survival and the proportion of the cohort progression free/ alive at 5 years for each fitted curve. For comparison, the tables include estimates of median and 5-year survival from other sources;

- the CABOSUN and COMPARZ trials.
- a cohort of patients from the IMDC database starting first-line treatment with sunitinib or pazopanib for metastatic RCC.[47]
- the committee's preferred model for the NICE appraisal of tivozanib (as reported in the published guidance).[22]

The COMPARZ and IMDC datasets and tivozanib model relate to patients with a mixed risk profile. Thus, estimates from these sources should be considered upper limits for survival for the intermediate/poor risk population in this appraisal.

Validity of fitted OS curves

For OS we have two sets of median survival estimates from CABOSUN, from the January 2017 and July 2017 cut-offs. The KM data used to fit the ITC models (clarification question B1) relates to the earlier cut-off. With respect to this dataset, several fitted models appear to underestimate median OS with cabozantinib, but for sunitinib median OS estimates were similar. For some of the fitted models, estimates of 5-year survival with sunitinib appear optimistic for the intermediate/poor risk population, as rates were similar in the IMDC cohort (21%).

Validity of fitted PFS curves

For sunitinib, most fitted models give estimates of median PFS that are similar to that in CABOSUN (5.3 months). Exceptions are the Weibull direct comparison; ITC exponential, Weibull and Gompertz; and first-order FPs. All the fitted models overestimated median PFS with cabozantinib with respect to CABOSUN (8.6 months). Median PFS estimates from the ITC models were slightly lower for pazopanib than for sunitinib; reflecting the small (but non-significant) PFS advantage for sunitinib in COMPARZ. As expected, all ITC models for sunitinib and pazopanib gave lower estimates of median PFS than the other sources. Five-year PFS with sunitinib and pazopanib was also lower for most fitted models than in the IMDC cohort, although sunitinib estimates from some FPs were similar to the 5% IMDC figure.

In summary, the ERG concludes that the company's preferred survival models (lognormal for the direct comparison and FP with P1=P2=-1 for the ITC) have reasonable face validity for sunitinib and pazopanib, with good measures of fit and similar median PFS as in the CABOSUN control arm. Both curves overestimate PFS for cabozantinib; yielding higher median PFS than in CABOSUN and a relatively large proportion of patients without disease progression at 5 years (5% and 10%). For OS, the company's preferred models (exponential direct comparison and FP with P1=P2=-1) also have a good fit for sunitinib and pazopanib and median survival is similar to that in the CABOSUN control arm (January 2017 data cut). For cabozantinib, the company's preferred models also give similar results to this dataset, but they overestimate median OS in relation to the most recent, July 2017 dataset. The plausibility of the company's survival extrapolations is unclear. 5-year survival with cabozantinib is estimated at 21% and 24% with the company's preferred direct and ITC models. To put this in

perspective, this is similar to 5-year OS in the IMDC cohort, who had a more favourable risk profile but were treated at first line with sunitinib or pazopanib.

Although we are critical of the apparent overestimation of PFS and OS for cabozantinib with the company's preferred methods, the other fitted models do not address these concerns.

4.3.8 Cost effectiveness results

Results from the company's economic model are presented in section B.3.7 of the CS. The ERG believes that these results include some errors in model inputs and calculations, as described in the previous section. For comparison, we reproduce the CS original results in Table 24 below and ERG corrected results in section 4.4.2.

For the company base case using the direct comparison from CABOSUN, an ICER of £37,793 per QALY gained is reported for cabozantinib versus sunitinib. Based on the company's preferred ITC model, sunitinib is dominated by pazopanib and the ICER for cabozantinib compared with pazopanib is £48,451. The pairwise ICER for cabozantinib compared with sunitinib in this model is £31,538.

Table 24 Company base-case results, deterministic (from CS Tables 60 and 61)

Drug	Costs (£)	QALYs	Life-years	ICER (£ per QALY gained)	
				Incremental analysis	Pairwise, cabozantinib vs. comparator
Direct comparison (CABOSUN)					
Sunitinib	[redacted]	[redacted]	[redacted]	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	37,793	37,793
ITC comparison (CABOSUN and COMPARZ)					
Pazopanib	[redacted]	[redacted]	[redacted]	-	48,451
Sunitinib	[redacted]	[redacted]	[redacted]	Dominated	31,538
Cabozantinib	[redacted]	[redacted]	[redacted]	48,451	-

The CS states that cabozantinib is an effective treatment for advanced RCC in treatment naïve patients when compared with sunitinib and pazopanib. No claims are made regarding cost-effectiveness and the company did not carry out any economic analysis for subgroups. The company's approach to handling uncertainty is discussed below.

4.3.9 Assessment of uncertainty

4.3.9.1 Probabilistic sensitivity analysis

The company PSA results are summarised in scatterplots, CEACs and tables of incremental cost per QALY gained (CS Figures 17 to 22: CS Tables 62 to 64). The PSA results are stable and similar to the deterministic results. The CS summarises the probabilistic results stating that there is a 66.1% probability (based on the CABOSUN study) or a 74.4% probability (based on the ITC result) of cabozantinib being cost-effective, relative to sunitinib, at a threshold willingness to pay of £50,000 per QALY gained. Relative to pazopanib, the CS quotes a 47.8% probability (based on the ITC result) of cabozantinib being cost-effective, at a threshold willingness to pay of £50,000 per QALY gained.

4.3.9.2 One-way sensitivity analyses

One-way sensitivity analyses are undertaken and reported in the CS. Model parameters are varied across a range to test the sensitivity of the ICERs to individual parameters or groups of parameters. The CS reports the input ranges and distributions for the model parameters in CS Table 65. The results are summarised in the tornado graphs in CS Figures 23 and 24. The company does not expressly justify the ranges used for the one-way sensitivity analysis. However, and most parameter ranges are based on observed values, such as 95% CI, and choice of distributions (CS Table 65) is reasonable. The tornado graphs only show parameters that make at least £1000 per QALY gained difference between the minimum and maximum limits. These include drug costs and discount rates for QALYs and costs, which have the biggest impact on cost-effectiveness. Other than these parameters, the key drivers of cost-effectiveness are relative dose intensity and utilities associated with the progression free state. However, we note that this analysis does not reflect uncertainties over the treatment effects on PFS, OS and TTD: structural uncertainty over the choice of survival curve analysis method; or uncertainty around the fitted parameters for those curves. The impact of these uncertainties is reflected in the PSA and scenario analyses.

4.3.9.3 Scenario analyses

The company explores a range of scenarios which are reported in the CS Table 66 and 67 Some of the company's scenario analyses were informed by expert opinion. Generally, the company appears to test scenarios using available data that was not used in the base case. The company found that the biggest source of uncertainty over cost-effectiveness was the choice of OS curve used in the model.

4.4 Additional work undertaken by the ERG

4.4.1 Description and justification of ERG analyses

Table 25 shows the corrections the ERG made to the company's model. Table 26 shows our preferred assumptions and scenarios, and Table 27 shows our approach to modelling treatment effects.

Table 25 ERG corrections to company model

Aspect of model	Problem	ERG correction
QALY calculations	1. Discounting and adjustment of utilities for age and sex are applied twice. 2. QALY loss for adverse events applied at first cycle is incorrectly adjusted for the duration of the cycle. 3. Accrual utility adjusted again for duration of cycle.	Columns AA to AF on Markov trace sheets ('E.Cabo.RCT' etc.) recoded
Health state utilities for scenario	4. Incorrect values for sunitinib TA scenario analysis in 'Resources' sheet (cells F231-M231).	Corrected in company response to clarification question.
Administration cost for nivolumab	5. Cost for administration not included in cost calculation. The cost cited on the User-Inputs page also differs to that for bevacizumab and temsirolimus.	Changed admin cost to £205 (cell I126 'User_Inputs') and added to weekly cost calculation for nivolumab (cell F97 'Variables').
Cost of end of life care at last cycle before death	6. Cost used in company base case only relates to hospital care. Costs for local authority funded social care, district nursing and GP visits excluded. Uprated using general price inflation (not health specific).	ERG estimated value of £7,961 for NHS and PSS perspective uprated from 2010/11 to 2016/17 using HCHS index.[52]
Scenario analyses	7. Scenario with PFS=OS=Gompertz used lognormal rather than Gompertz distribution for PFS. 8. Health resource (UK clinicians) scenario gives base case result.	'ScenarioAnalysis' cell N33 changed to 5 (Gompertz). Control in cell F154 on the 'User_Inputs' sheet linked to v.vHealthResource.Input.
TTD on curve data sheet and graph	9. TTD curve for pazopanib defined in relation to pazopanib PFS curve. Gives wrong summary statistics and TTD curve on 'User_Inputs' sheet.	Deleted 'Curve data' sheet and replaced summary statistic calculations and figures linked to 'TPs_CABOSUN' etc. sheets

Table 26 ERG preferred assumptions and scenarios

	Preferred assumptions	Scenarios	Reason for analysis
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=80 (July 2017 CABOSUN update). OS assumed equal for pazopanib and sunitinib, based on 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 >=5% incidence	Range of disutilites, 8 week duration and >=2%	Again, we follow the company approach, but conduct additional analyses to test the sensitivity of the model to adverse events.

	Preferred assumptions	Scenarios	Reason for analysis
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	years	55/75 years	Exploratory: to assess applicability to the UK RCC population

Table 27 ERG approach to modelling treatment effects

	Company base case (scenarios)	Comments	ERG preferred assumptions
OS curves	**Direct**: Exponential (Weibull & Gompertz) **ITC**: FP model with P1=P2=-1 (exponential; Weibull; Gompertz; and FP P1=-0.5, P2=0 & P1=-1, P2=0)	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. Given these reservations, the exponential, Weibull and Gompertz are reasonable for the direct analysis. For the ITC, the exponential and FP P1=P2=-1 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.	Simple indirect comparison assuming: • Sunitinib OS curve based on company's exponential fit to CABOSUN; • Cabozantinib calculated from sunitinib curve and HR form July 2017 CABOSUN results; • Pazopanib curve assumed equal to sunitinib (based on COMPARZ results).
PFS curves	**Direct**: lognormal (Exponential, Weibull & Gompertz) **ITC**: FP P1=P2=-1 (exponential, Weibull and Gompertz)	CABOSUN PFS analysis is more mature. ITC is subject to uncertainty due to differences in trial populations, unclear if similarity assumption is met. 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.	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.
TTD curves	**Direct**: lognormal (exponential, Weibull, Gompertz & gamma).	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.	Lognormal for base case, and all other distributions in scenario analysis.

4.4.2 Results of ERG analyses

All analyses in this report reflect agreed PAS discounts for cabozantinib and pazopanib, and the free first cycle for sunitinib. However, they exclude PAS discounts for subsequent treatments. PAS discounts are in place for cabozantinib, axitinib, nivolumab, everolimus and lenvatinib. We replicate the tables below with PAS discounts in a confidential addendum to this report.

4.4.2.1 ERG corrections to company analyses

Corrections to the QALY calculations (points 1-3 in Table 25) increase the QALY estimates for all treatments. This increases the incremental QALY gains for cabozantinib, hence reducing ICERs. For example, based on the direct comparison with sunitinib, incremental QALYs increase from 0.401 in the original company base case to 0.471 in our corrected analysis, which reduces the ICER from £37,392 to £32,340 per QALY gained. Corrections to the costs of nivolumab and end of life care further reduce the ICER estimates: e.g. to £31,956 for the direct comparison of cabozantinib with sunitinib. The cost-effectiveness results from the ERG corrections to the company's base case analyses are shown in Table 28. These show that sunitinib is dominated by pazopanib, which yields more QALYs at a lower cost. The ICER for cabozantinib compared with pazopanib is £40,757 per QALY gained. Compared with sunitinib, cabozantinib has an ICER of £31,956 per QALY gained based on the direct comparison from CABOSUN data, and £26,182 per QALY gained based on the company's preferred indirect comparison using CABOSUN and COMPARZ data.

Table 28 Cost-effectiveness: Company base-case analyses (ERG corrected)

Drug	Costs (£)	QALYs	Life-years	ICER (£ per QALY gained)	
				Incremental analysis	Pairwise, cabozantinib vs. comparator
Direct comparison (CABOSUN)					
Sunitinib	[redacted]	[redacted]	[redacted]	-	31,956
Cabozantinib	[redacted]	[redacted]	[redacted]	31,956	-
ITC (CABOSUN and COMPARZ)					
Pazopanib	[redacted]	[redacted]	[redacted]	-	40,757
Sunitinib	[redacted]	[redacted]	[redacted]	Dominated	26,182
Cabozantinib	[redacted]	[redacted]	[redacted]	40,757	-

Results from the probabilistic sensitivity analysis (PSA) are similar. The extent of uncertainty around the incremental costs and QALYs for cabozantinib compared with sunitinib and pazopanib is illustrated in the scatterplots in Figure 20 (for the company's ITC base case).

Based on 1,000 PSA iterations, there is an estimated probability that cabozantinib is cost-effective compared with pazopanib is 28% at a cost effectiveness threshold of £30,000 per QALY gained and 57% at a threshold of £50,000 per QALY gained.

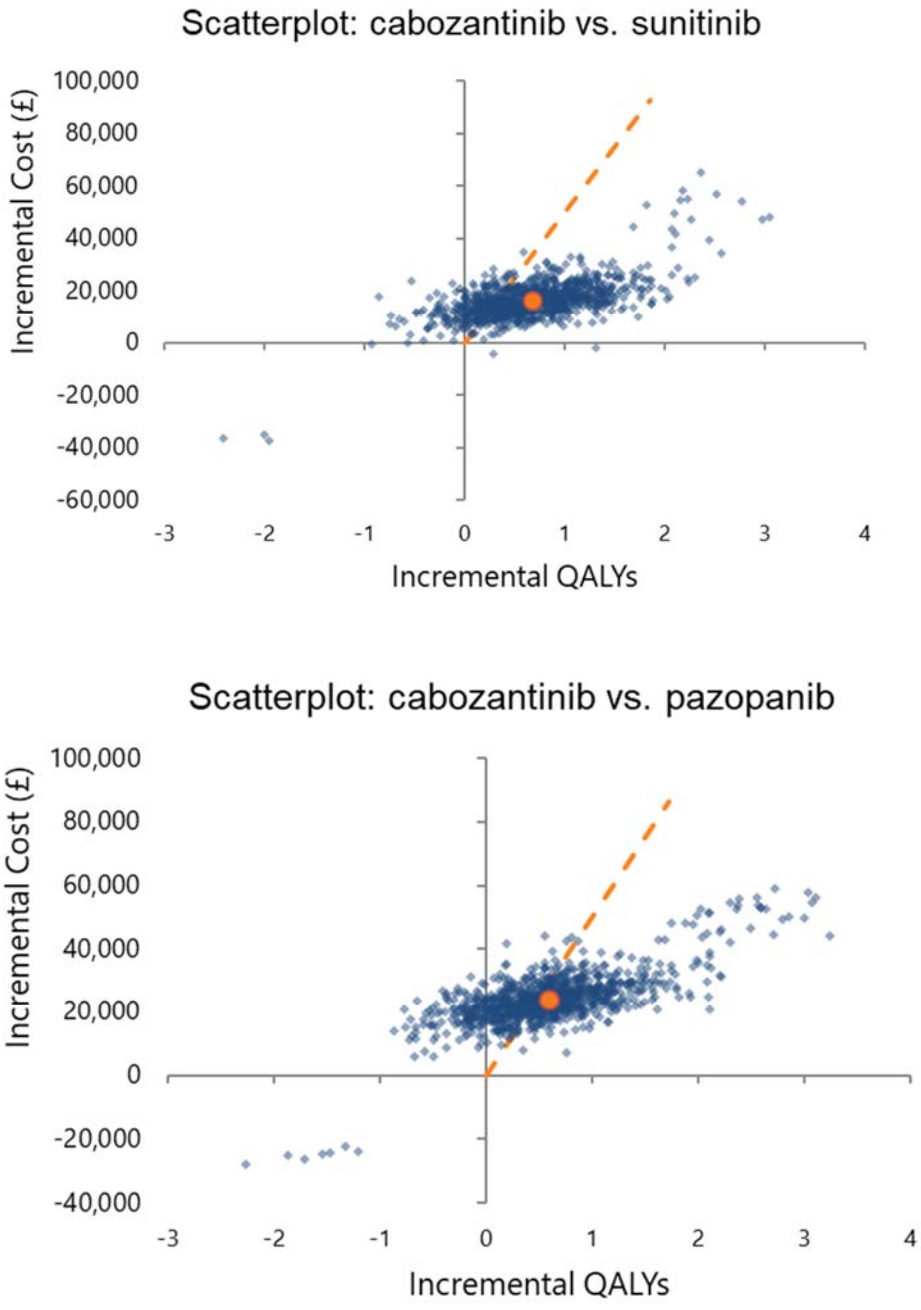


Figure 20 CE scatterplots, company ITC base case (ERG corrected)

Results from the one-way deterministic sensitivity analyses on the ERG-corrected version of the company's ITC base case are illustrated in Figure 21 and Figure 22. These suggest that the cost and relative dose intensity of the treatment and comparator as well as other cost parameters are the key drivers of cost-effectiveness. However, this is misleading, as effectiveness parameters are not included in this analysis.

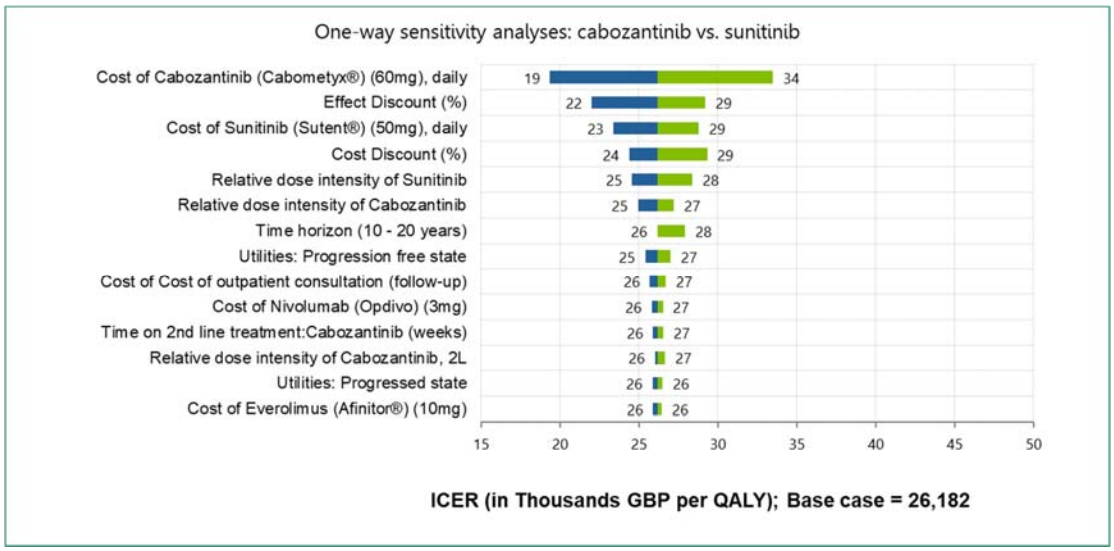


Figure 21 Tornado diagram: Company ITC base case (ERG corrected)

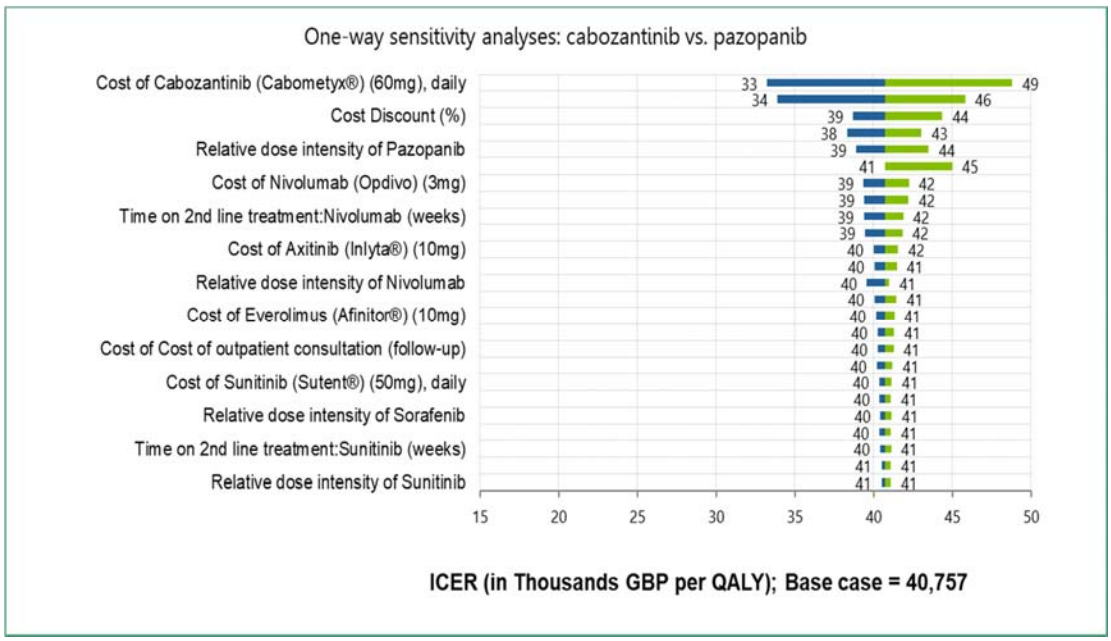


Figure 22 Tornado diagram: Company ITC base case (ERG corrected)

The impact of key uncertainties over model assumptions and data sources of data is shown in Table 29, Table 30 and Table 31 below. The model is most sensitive to assumptions and methods of fitting the OS curves. The model is also sensitive to a very short time horizon.

Table 29 Scenario analysis: Company direct base case (ERG corrected) vs. sunitinib

Scenario		Total cost (£)		Total QALYs		ICER (£ per QALY)
		Cabo	Suni	Cabo	Suni	
Company direct base case		[redacted]	[redacted]	[redacted]	[redacted]	**31,956**
Time horizon	10 years	[redacted]	[redacted]	[redacted]	[redacted]	33,216
	5 years	[redacted]	[redacted]	[redacted]	[redacted]	40,719
PFS curves	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	30,414
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	29,247
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	31,562
	Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	31,749
	Gamma	[redacted]	[redacted]	[redacted]	[redacted]	30,671
OS curves	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	41,669
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	30,226
	Lognormal	[redacted]	[redacted]	[redacted]	[redacted]	38,946
	Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	47,576
	Gamma	[redacted]	[redacted]	[redacted]	[redacted]	20,841
TTD curves	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	26,586
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	26,596
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	28,978
	Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	33,022
	Gamma	[redacted]	[redacted]	[redacted]	[redacted]	29,879
Utility source	PF and PD (Swinburn)	[redacted]	[redacted]	[redacted]	[redacted]	27,461
	PF and PD (Pazo TA215)	[redacted]	[redacted]	[redacted]	[redacted]	32,912
	PF and PD (Suni TA169)	[redacted]	[redacted]	[redacted]	[redacted]	29,779
	TEAE (METEOR)	[redacted]	[redacted]	[redacted]	[redacted]	31,893
Costs	Comprehensive blood test	[redacted]	[redacted]	[redacted]	[redacted]	32,266
	Management (UK clinician)	[redacted]	[redacted]	[redacted]	[redacted]	30,595
	Second-line (UK practice)	[redacted]	[redacted]	[redacted]	[redacted]	34,081
	No infusion wastage	[redacted]	[redacted]	[redacted]	[redacted]	32,099
	No end of life cost	[redacted]	[redacted]	[redacted]	[redacted]	32,349

Table 30 Scenario analysis: Company ITC base case (ERG corrected), vs. sunitinib

Scenario		Total cost (£)		Total QALYs		ICER (£ per QALY)
		Cabo	Suni	Cabo	Suni	
Company ITC base case		[redacted]	[redacted]	[redacted]	[redacted]	**26,182**
Time horizon	10 years	[redacted]	[redacted]	[redacted]	[redacted]	27,912
	5 years	[redacted]	[redacted]	[redacted]	[redacted]	35,488
PFS curves	FP P1=-0.5 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	25,795
	FP P1=-1 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	25,818
	ITC RE exponential	[redacted]	[redacted]	[redacted]	[redacted]	28,909
	ITC RE Weibull	[redacted]	[redacted]	[redacted]	[redacted]	28,551
	ITC RE Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	29,043
	ITC RE Lognormal	[redacted]	[redacted]	[redacted]	[redacted]	30,094
	ITC RE Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	29,700
OS curves	FP P1=-0.5 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	55,215
	FP P1=-1 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	33,356
	ITC RE Exponential	[redacted]	[redacted]	[redacted]	[redacted]	30,094
	ITC RE Weibull	[redacted]	[redacted]	[redacted]	[redacted]	38,252
	ITC RE Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	23,445
	ITC RE Lognormal	[redacted]	[redacted]	[redacted]	[redacted]	45,415
	ITC RE Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	49,983
TTD curves	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	21,816
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	21,826
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	23,760
	Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	27,702
	Gamma	[redacted]	[redacted]	[redacted]	[redacted]	24,475
Utility source	PF and PD (Swinburn)	[redacted]	[redacted]	[redacted]	[redacted]	21,332
	PF and PD (Pazo TA215)	[redacted]	[redacted]	[redacted]	[redacted]	26,787
	PF and PD (Suni TA169)	[redacted]	[redacted]	[redacted]	[redacted]	24,431
	TEAE (METEOR)	[redacted]	[redacted]	[redacted]	[redacted]	26,141
Costs	Comprehensive blood test	[redacted]	[redacted]	[redacted]	[redacted]	26,488
	Management (UK clinician)	[redacted]	[redacted]	[redacted]	[redacted]	24,926
	Second-line (UK practice)	[redacted]	[redacted]	[redacted]	[redacted]	28,425
	No infusion wastage	[redacted]	[redacted]	[redacted]	[redacted]	26,262
	No end of life cost	[redacted]	[redacted]	[redacted]	[redacted]	26,585

Table 31 Scenario analysis: Company ITC base case (ERG corrected), vs. pazopanib

Scenario		Total cost (£)		Total QALYs		ICER (£ per QALY)
		Cabo	Pazo	Cabo	Pazo	
Company ITC base case		[redacted]	[redacted]	[redacted]	[redacted]	**40,757**
Time horizon	10 years	[redacted]	[redacted]	[redacted]	[redacted]	45,001
	5 years	[redacted]	[redacted]	[redacted]	[redacted]	64,841
PFS curves	FP P1=-0.5 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	39,653
	FP P1=-1 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	39,733
	ITC RE exponential	[redacted]	[redacted]	[redacted]	[redacted]	50,540
	ITC RE Weibull	[redacted]	[redacted]	[redacted]	[redacted]	50,591
	ITC RE Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	49,206
	ITC RE Lognormal	[redacted]	[redacted]	[redacted]	[redacted]	51,910
	ITC RE Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	50,037
OS curves	FP P1=-0.5 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	74,858
	FP P1=-1 P2=0	[redacted]	[redacted]	[redacted]	[redacted]	49,973
	ITC RE Exponential	[redacted]	[redacted]	[redacted]	[redacted]	51,910
	ITC RE Weibull	[redacted]	[redacted]	[redacted]	[redacted]	65,942
	ITC RE Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	37,788
	ITC RE Lognormal	[redacted]	[redacted]	[redacted]	[redacted]	78,883
	ITC RE Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	86,300
TTD curves	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	36,236
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	36,188
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	38,277
	Loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	42,564
	Gamma	[redacted]	[redacted]	[redacted]	[redacted]	38,826
Utility source	PF and PD (Swinburn)	[redacted]	[redacted]	[redacted]	[redacted]	31,471
	PF and PD (Pazo TA215)	[redacted]	[redacted]	[redacted]	[redacted]	41,419
	PF and PD (Suni TA169)	[redacted]	[redacted]	[redacted]	[redacted]	38,073
	TEAE (METEOR)	[redacted]	[redacted]	[redacted]	[redacted]	40,578
Costs	Comprehensive blood test	[redacted]	[redacted]	[redacted]	[redacted]	41,060
	Management (UK clinician)	[redacted]	[redacted]	[redacted]	[redacted]	39,603
	Second-line (UK practice)	[redacted]	[redacted]	[redacted]	[redacted]	26,736
	No infusion wastage	[redacted]	[redacted]	[redacted]	[redacted]	39,979
	No end of life cost	[redacted]	[redacted]	[redacted]	[redacted]	41,159

4.4.2.2 ERG preferred analysis

Results based on the ERG preferred assumptions are shown in Table 32. As in the company analyses, sunitinib is dominated as pazopanib is less expensive and no less effective. Compared with pazopanib, cabozantinib has an ICER of £65,743 per QALY gained. The ICER for cabozantinib is £41,465 in comparison with sunitinib. By assumption, life expectancy is the same for pazopanib and sunitinib in this analysis and there is a small difference in mean progression-free life years between these comparators. Cabozantinib has a modest survival advantage and a larger effect on progression free-survival. We believe that these estimates appropriately reflect the evidence from the CABOSUN and COMPARZ trials.

Table 32 Cost-effectiveness: ERG preferred assumptions

Drug	Costs (£)	QALYs	Life-years	PF life years	ICER (£ per QALY gained)	
					Incremental analysis	Pairwise, cabozantinib vs. comparator
Pazopanib	[redacted]	[redacted]	[redacted]	[redacted]	-	65,743
Sunitinib	[redacted]	[redacted]	[redacted]	[redacted]	-	41,465
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	65,743	-

Table 33 and Table 34 summarise scenario analyses around our preferred set of assumptions for the comparison of cabozantinib versus pazopanib and sunitinib respectively. Generally, the results are robust, with the pairwise ICERs remaining above £30,000 per QALY gained for all scenarios tested. The ICERs were most sensitive to the assumption that cabozantinib has no relative effect on survival compared with sunitinib or pazopanib. This illustrates that the results are very largely driven by the effect on OS, as estimated from the CABOSUN trial.

Table 33 Scenario analysis: ERG preferred assumptions, vs. pazopanib

Scenario		Total cost (£)		Total QALY		ICER (£)
		Cabo.	Pazo.	Cabo.	Pazo.	
ERG preferred assumptions		[redacted]	[redacted]	[redacted]	[redacted]	**65,743**
Time horizon	5 years	[redacted]	[redacted]	[redacted]	[redacted]	79,127
	10 years	[redacted]	[redacted]	[redacted]	[redacted]	66,783
Persistence of OS/ PFS effect	10 years	[redacted]	[redacted]	[redacted]	[redacted]	58,890
	20 years	[redacted]	[redacted]	[redacted]	[redacted]	57,879
CABOSUN OS curves	HR = 0.74 (Jan 2017)	[redacted]	[redacted]	[redacted]	[redacted]	52,778
	No effect on OS	[redacted]	[redacted]	[redacted]	[redacted]	372,866
CABOSUN PFS curves	Separate exponential	[redacted]	[redacted]	[redacted]	[redacted]	64,913
	Separate Weibull	[redacted]	[redacted]	[redacted]	[redacted]	64,192
	Separate Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	64,880
TTD curves	Separate exponential	[redacted]	[redacted]	[redacted]	[redacted]	59,908
	Separate Weibull	[redacted]	[redacted]	[redacted]	[redacted]	59,836
	Separate Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	63,012
	Separate loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	65,638
	Separate gamma	[redacted]	[redacted]	[redacted]	[redacted]	64,092
Utility values	Swinburn	[redacted]	[redacted]	[redacted]	[redacted]	47,616
	Pazo NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	66,246
	Suni NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	61,500
TEAE disutility	METEOR (-0.05)	[redacted]	[redacted]	[redacted]	[redacted]	65,224
	Higher disutility (-0.4)	[redacted]	[redacted]	[redacted]	[redacted]	66,436
	Include if >= 2%	[redacted]	[redacted]	[redacted]	[redacted]	64,863
	Duration: 8 weeks	[redacted]	[redacted]	[redacted]	[redacted]	66,468
Drug costs (first line)	Dose intensities 86%	[redacted]	[redacted]	[redacted]	[redacted]	58,517
	Does intensities 100%	[redacted]	[redacted]	[redacted]	[redacted]	65,739
Drug costs (second line)	% use (Company)	[redacted]	[redacted]	[redacted]	[redacted]	41,936
	% use (ERG 1)	[redacted]	[redacted]	[redacted]	[redacted]	45,980
	% use (ERG 2)	[redacted]	[redacted]	[redacted]	[redacted]	44,374
Other costs	Blood test (£20)	[redacted]	[redacted]	[redacted]	[redacted]	66,039
	Follow up (UK clinician)	[redacted]	[redacted]	[redacted]	[redacted]	64,738
	No end of life cost	[redacted]	[redacted]	[redacted]	[redacted]	66,106
Age of cohort	55 years	[redacted]	[redacted]	[redacted]	[redacted]	65,567
	75 years	[redacted]	[redacted]	[redacted]	[redacted]	66,061

Table 34 Scenario analysis: ERG preferred assumptions, vs. sunitinib

Scenario		Total cost (£)		Total QALY		ICER (£)
		Cabo.	Suni.	Cabo.	Suni.	
ERG preferred assumptions		[redacted]	[redacted]	[redacted]	[redacted]	**41,465**
Time horizon	5 years	[redacted]	[redacted]	[redacted]	[redacted]	46,564
	10 years	[redacted]	[redacted]	[redacted]	[redacted]	41,839
Persistence of OS/ PFS effect	10 years	[redacted]	[redacted]	[redacted]	[redacted]	37,716
	20 years	[redacted]	[redacted]	[redacted]	[redacted]	37,170
CABOSUN OS curves	Exponential, HR 0.74	[redacted]	[redacted]	[redacted]	[redacted]	34,202
	No effect on OS	[redacted]	[redacted]	[redacted]	[redacted]	204,789
CABOSUN PFS curves	Separate exponential	[redacted]	[redacted]	[redacted]	[redacted]	39,904
	Separate Weibull	[redacted]	[redacted]	[redacted]	[redacted]	38,871
	Separate Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	40,107
TTD curves	Separate exponential	[redacted]	[redacted]	[redacted]	[redacted]	35,219
	Separate Weibull	[redacted]	[redacted]	[redacted]	[redacted]	35,237
	Separate Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	38,267
	Separate loglogistic	[redacted]	[redacted]	[redacted]	[redacted]	41,428
	Separate gamma	[redacted]	[redacted]	[redacted]	[redacted]	39,696
Utility values	Swinburn	[redacted]	[redacted]	[redacted]	[redacted]	30,089
	Pazo NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	41,780
	Suni NICE STA	[redacted]	[redacted]	[redacted]	[redacted]	38,805
TEAE disutility	METEOR (-0.05)	[redacted]	[redacted]	[redacted]	[redacted]	41,346
	Higher disutility (-0.4)	[redacted]	[redacted]	[redacted]	[redacted]	41,621
	Include if >= 2%	[redacted]	[redacted]	[redacted]	[redacted]	41,026
	Duration: 8 weeks	[redacted]	[redacted]	[redacted]	[redacted]	41,628
Drug costs (first line)	Dose intensities 86%	[redacted]	[redacted]	[redacted]	[redacted]	34,713
	Does intensities 100%	[redacted]	[redacted]	[redacted]	[redacted]	42,158
Drug costs (second line)	% use (Company)	[redacted]	[redacted]	[redacted]	[redacted]	43,856
	% use (ERG 1)	[redacted]	[redacted]	[redacted]	[redacted]	47,872
	% use (ERG 2)	[redacted]	[redacted]	[redacted]	[redacted]	46,276
Other costs	Blood test (£20)	[redacted]	[redacted]	[redacted]	[redacted]	41,759
	Follow up (UK clinician)	[redacted]	[redacted]	[redacted]	[redacted]	40,466
	No end of life cost	[redacted]	[redacted]	[redacted]	[redacted]	41,825
Age of cohort	55 years	[redacted]	[redacted]	[redacted]	[redacted]	41,354
	75 years	[redacted]	[redacted]	[redacted]	[redacted]	41,664

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)

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	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.
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	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 ([redacted] 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.

6 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

(MET, AXL), as well as tumour angiogenesis (VEGF). It is stated that by targeting MET and AXL receptors in addition to VEGFR, cabozantinib may provide additional anticancer efficacy over the more selective, existing anti-VEGFR agents (B.2.12).

Expert clinical advice to the ERG suggests that all of the currently available drugs inhibit VEGF, which is thought to be the main mechanism of action in RCC. Cabozantinib is not the only drug therapy that targets other pathways. However, it is not yet clear how important these other pathways are in drug efficacy.

The ERG notes that in the previous NICE appraisal of cabozantinib for previously treated RCC (TA 463)[14] it was accepted that cabozantinib would likely have additional benefits for some patients due to its multi-targeted approach, and could therefore be considered innovative. However, cabozantinib was not considered to reflect a 'step change' in treatment (The ERG infers that this consideration is within the context of previously treated RCC patients, not necessarily within the context of untreated RCC).

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

The results of the CABOSUN trial show a statistically significant effect on PFS, the primary outcome, with a median PFS of 8.6 months (95% CI 6.8, 14.0) for cabozantinib and 5.3 months (95% CI 3.0, 8.2) for sunitinib (p=0.0008). The median difference of 3.3 months favoured cabozantinib. The confidence intervals around the PFS estimates are reasonably narrow indicating greater certainty in the estimates. It is important to put these results into context of the results of other trials of first line drug therapies in RCC. Rini and Vogelzang[63] discussed the results of the CABOSUN trial and noted that the median PFS of 5.6 months for the sunitinib arm of the CABOSUN trial was lower than that achieved in previous clinical trials. Specifically, in the phase III registration trial for sunitinib,[64] the median PFS for patients in the intermediate risk subgroup was 10.6 months. The ERG notes that this trial had a slightly lower percentage of patients with bone metastases and lower percentage of patients with prior nephrectomy than CABOSUN, which suggests slightly more favourable prognostic characteristics. Nonetheless, it can be considered an informative benchmark for PFS. The ERG notes that in a recently published phase III RCT comparing nivolumab in combination with ipilimumab versus sunitinib

in previously untreated clear-cell advanced RCC, median PFS for the sunitinib arm in the intermediate/poor risk subgroup was 8.4 months. Of note, the statistical power calculation in the CABOSUN trial assumed a median PFS of 8 months for sunitinib. This is 2.7 months higher than the median PFS achieved. The CS does not comment on this.

The CS cites a registry study of 1189 previously untreated poor and intermediate risk patients receiving targeted therapies (among whom sunitinib was the most common treatment), which reported a PFS of 5.6 months.[65] The CS suggests this is consistent with the CABOSUN results. However, Rini and Vogelzang[63] note that this data set included patients with non–clear cell histology (12%), patients with sarcomatoid histology (10%), and patients who received sorafenib, temsirolimus, or everolimus (21%). They suggest that these features might be expected to result in a lower PFS than would be expected in practice and the benchmark of 5.6 months isn't necessarily realistic.

Choueiri et al[66] (the CABOSUN trial investigators) responded to Rini and Vogelzang[63] that the CABOSUN trial included patients with high rates of poor prognostic clinical factors, which distinguishes it from other contemporary trials of untreated patients with metastatic RCC. They note that PFS was also shorter in a retrospective community setting study of sunitinib (7.5 months) in 134 patients.[67] They describe this as an 'all comer' population, but don't define what this means. The ERG infers that it is likely to mean a population representative of community practice. Choueiri et al[66] state that the cooperative group setting (which they imply is relevant to the CABOSUN trial) is more akin to community practice. The ERG considers that this is a plausible explanation for differences between the sunitinib PFS results of the trial compared to other trials.

Another finding from the CABOSUN trial was that there was a statistically significant difference in the ORR between cabozantinib and sunitinib, favouring cabozantinib. All responses were classed as a 'confirmed partial response', and there were no confirmed complete responders in either study group. Expert clinical advice to the ERG suggests that a complete response would not necessarily be expected in an intermediate or poor risk patient group, and that genuine complete responders to these agents would be relatively unusual.

7.2 Summary of cost effectiveness issues

In the company's analysis the direct comparison of cabozantinib with sunitinib, based on extrapolation of OS, PFS and TTD curves from CABOSUN, gave an ICER of £37,793 per QALY gained. The indirect comparison, with OS and PFS extrapolations based on the fractional polynomial ITC, gave an ICER of £31,538 for cabozantinib compared with sunitinib and £48,451 for cabozantinib compared with pazopanib. In this analysis, pazopanib had lower a lower mean cost and higher mean QALYs than sunitinib: sunitinib is dominated. The company's analysis of uncertainty identifies the OS curves and the cost of cabozantinib as the main drivers of cost-effectiveness.

The ERG identified and corrected some errors and inconsistencies in the company's submitted model the most significant of which was a coding error in QALY calculations that had the effect of underestimating QALYs for each treatment. This resulted in lower ICERs for the company's base cases: £31,956 per QALY for the direct comparison of cabozantinib with sunitinib; and for the ITC analysis, £40,757 for cabozantinib compared with pazopanib and £26,182 compared with sunitinib.

The ERG identified a number of uncertainties in the company's model and identified an alternative set of assumptions and input parameters relating to the method of fitting the OS curves, the time horizon and duration of effects, and health state utilities, adverse effects and costs.

The ERG-preferred analyses gave higher ICER estimates: £65,742 for cabozantinib compared with pazopanib and £41,465 compared with sunitinib. As in the company base case, we estimate that sunitinib is dominated by pazopanib due to its higher cost and similar effectiveness. However, this result was sensitive to some cost and resource use assumptions. By assumption, our preferred analysis gave the same life expectancy with sunitinib as with pazopanib, yielding very similar QALY estimates. Cabozantinib has a modest survival advantage and a larger effect on progression free survival and hence QALYs. We believe that these results appropriately reflect evidence from CABOSUN and COMPARZ. The results were generally robust, with the ICERs remaining above £30,000 per QALY gained for all of the scenarios that we tested.

The above analyses include existing PAS discounts for cabozantinib, sunitinib and pazopanib for first-line treatments. However, they exclude these arrangements and other existing PAS discounts for subsequent treatment after failure of first line treatment. We present results for the ERG-corrected company base case and scenarios and for ERG additional analysis in a confidential addendum to this report.

8 REFERENCES

1. Escudier B, Porta C, Schmidinger M, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* 2016;27(suppl_5):v58-v68. doi: 10.1093/annonc/mdw328
2. Cancer Research UK. Kidney cancer [updated 14/02/2018.
3. Kim DY, Wood CG, Karam JA. Treating the two extremes in renal cell carcinoma: management of small renal masses and cytoreductive nephrectomy in metastatic disease. *American Society of Clinical Oncology educational book American Society of Clinical Oncology Meeting* 2014:e214-21. doi: 10.14694/EdBook_AM.2014.34.e214 [published Online First: 2014/05/27]
4. Ljungberg B, Hanbury DC, Kuczyk MA, et al. Renal Cell Carcinoma Guideline. *European urology* 2007;51(6):1502-10. doi: 10.1016/j.eururo.2007.03.035
5. National Cancer Registration and Analysis Service. TNM stage group by CCG by tumor type for 10+3 tumour types, 2015 [Available from: http://www.ncin.org.uk/publications/survival_by_stage accessed 21/02/2018.
6. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology: Kidney Cancer, version 2.2018. *https://wwwnccnorg/professionals/physician_gls/defaultaspx* 2017
7. Heng DY, Xie W, Regan MM, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. *The Lancet Oncology* 2013;14(2):141-8. doi: 10.1016/s1470-2045(12)70559-4 [published Online First: 2013/01/15]
8. Motzer RJ, Bacik J, Schwartz LH, et al. Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2004;22(3):454-63. doi: 10.1200/jco.2004.06.132 [published Online First: 2004/01/31]

9. Harding G, Cella D, Robinson D, Jr., et al. Symptom burden among patients with renal cell carcinoma (RCC): content for a symptom index. *Health and quality of life outcomes* 2007;5:34. doi: 10.1186/1477-7525-5-34 [published Online First: 2007/06/16]
10. Cella D. Beyond traditional outcomes: improving quality of life in patients with renal cell carcinoma. *The oncologist* 2011;16 Suppl 2:23-31. doi: 10.1634/theoncologist.2011-S2-23 [published Online First: 2011/03/05]
11. Gupta K, Miller JD, Li JZ, et al. Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): a literature review. *Cancer treatment reviews* 2008;34(3):193-205. doi: 10.1016/j.ctrv.2007.12.001 [published Online First: 2008/03/04]
12. National Institute for Health and Care Excellence. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma. Technology appraisal guidance (TA169). London, 2009.
13. National Institute for Health and Care Excellence. Pazopanib for the first-line treatment of advanced renal cell carcinoma. Technology appraisal guidance (TA215). London, 2013.
14. National Institute for Health and Care Excellence. Cabozantinib for previously treated advanced renal cell carcinoma. Technology appraisal guidance (TA463). London: NICE, 2017.
15. Swinburn P, Lloyd A, Nathan P, et al. Elicitation of health state utilities in metastatic renal cell carcinoma. *Current medical research and opinion* 2010;26(5):1091-6. doi: 10.1185/03007991003712258 [published Online First: 2010/03/17]
16. Office for National Statistics. Cancer registration statistics, England: first release, 2016 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2016 accessed 21/02/2018.
17. Chow WH, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. *Nature reviews Urology* 2010;7(5):245-57. doi: 10.1038/nrurol.2010.46 [published Online First: 2010/05/08]
18. Verine J, Pluvinage A, Bousquet G, et al. Hereditary renal cancer syndromes: an update of a systematic review. *European urology* 2010;58(5):701-10. doi: 10.1016/j.eururo.2010.08.031 [published Online First: 2010/09/08]
19. Koneru R, Hotte SJ. Role of cytokine therapy for renal cell carcinoma in the era of targeted agents. *Current oncology (Toronto, Ont)* 2009;16 Suppl 1:S40-4. [published Online First: 2009/05/30]

20. National Institute for Health and Care 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. Technology appraisal guidance (TA178). London, 2009.
21. National Institute for Health and Care Excellence. Renal cancer overview. https://pathwaysniceorguk/pathways/renal-cancer.
22. National Institute for Health and Care Excellence. Tivozanib for treating advanced renal cell carcinoma (TA512). London, 2018.
23. Powles T, Albiges L, Staehler M, et al. Updated European Association of Urology Guidelines Recommendations for the Treatment of First-line Metastatic Clear Cell Renal Cancer. *European urology* 2017 doi: 10.1016/j.eururo.2017.11.016 [published Online First: 2017/12/11]
24. Choueiri TK, Halabi S, Sanford BL, et al. Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Patients With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2017;35(6):591-97. doi: 10.1200/jco.2016.70.7398 [published Online First: 2017/02/16]
25. Choueiri TK, Hessel C, Halabi S, et al. Progression-free survival (PFS) by independent review and updated overall survival (OS) results from Alliance A031203 trial (CABOSUN): Cabozantinib versus sunitinib as initial targeted therapy for patients (pts) with metastatic renal cell carcinoma (mRCC). *Annals of Oncology* 2017;28:1.
26. Cabozantinib versus sunitinib for previously untreated patients with advanced renal cell carcinoma (RCC) of intermediate or poor risk: Subgroup analysis of progression-free survival (PFS) and objective response rate (ORR) in the Alliance A031203 CABOSUN trial. ASCO GenitoUrinary Cancer Symposium 2018 February 10, 2018; San Francisco, CA.
27. Exelixis Inc. Clinical Study Report: A031203 Randomized Phase II Study Comparing Cabozantinib (NSC #761968 and IND #116059) with Commercially Supplied Sunitinib in Subjects with Previously Untreated Locally Advanced or Metastatic Renal Cell Carcinoma. San Francisco, 2017.
28. Choueiri TK, Halabi S, Sanford B, et al. CABOzantinib versus SUNitinib (CABOSUN) as initial targeted therapy for patients with metastatic renal cell carcinoma (mRCC) of poor and intermediate risk groups: Results from ALLIANCE A031203 trial. *Annals of Oncology* 2016;27(suppl_6):LBA30_PR-LBA30_PR. doi: 10.1093/annonc/mdw435.23

29. Chopra M. Annual Congress of the European Society for Medical Oncology (ESMO): Copenhagen, Denmark; 7–11 October 2016, 2016.
30. Choueiri TK, Hessel C, Halbi S, et al. Progression-free survival by independent review and overall survival update for the Alliance A031203 CABOSUN trial of cabozantinib vs sunitinib in metastatic renal cell carcinoma. European Society for Medical Oncology Congress. Madrid, Spain, 2017.
31. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. York Publishing Services Ltd., 2009.
32. Pond GR. Statistical issues in the use of dynamic allocation methods for balancing baseline covariates. *British Journal of Cancer* 2011;104(11):1711-15. doi: 10.1038/bjc.2011.157
33. Wilson MK, Collyar D, Chingos DT, et al. Outcomes and endpoints in cancer trials: bridging the divide. *The Lancet Oncology* 2015;16(1):e43-e52. doi: 10.1016/S1470-2045(14)70380-8
34. Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus Sunitinib in Metastatic Renal-Cell Carcinoma. *New England Journal of Medicine* 2013;369(8):722-31. doi: 10.1056/NEJMoa1303989
35. Motzer RJ, Hutson TE, McCann L, et al. Overall Survival in Renal-Cell Carcinoma with Pazopanib versus Sunitinib. *New England Journal of Medicine* 2014;370(18):1769-70. doi: 10.1056/NEJMc1400731
36. Ouwens MJ, Philips Z, Jansen JP. Network meta-analysis of parametric survival curves. *Research synthesis methods* 2010;1(3-4):258-71. doi: 10.1002/jrsm.25 [published Online First: 2010/07/01]
37. ECOG-ACRIN Cancer Research Group. ECOG Performance Status 2018 [Available from: http://ecog-acrin.org/resources/ecog-performance-status accessed 13/3/18.
38. GlaxoSmithKline. Study VEG108844, a study of pazopanib versus sunitinib in the treatment of subjects with locally advanced and/or metastatic renal cell carcinoma.: GlaxoSmithKline, 2013.
39. GlaxoSmithKline. Study VEG108844, a study of pazopanib versus sunitinib in the treatment of subjects with locally advanced and/or metastatic renal cell carcinoma: overall survival update. Report dated 13 Mar 2014.: GlaxoSmithKline, 2014.
40. Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Medical Research Methodology* 2011;11(1):61. doi: 10.1186/1471-2288-11-61

41. National Institute for Health and Care Excellence. Fulvestrant for the treatment of locally advanced or metastatic breast cancer. Technolgy appraisal guidance (TA239). London: NICE, 2011.
42. National Institute for Health and Care Excellence. Fulvestrant for untreated locally advanced or metastatic oestrogen-receptor positive breast cancer. Technolgy appraisal guidance (TA503). London: NICE, 2018.
43. MRC Biostatistics Unit. DIC: Deviance Information Criteria 2017 [Available from: https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-dic/ accessed 12/3/18.
44. Gelman A, Shirley K. Chapter 6: Inference from simulations and monitoring convergence. In: Brooks S, Gelman A, Jones G, et al., eds. Handbook of Markov Chain Monte Carlo. Boca Raton, Florida: Chapman and Hall 2001.
45. Ara R, Brazier J. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. *Value In Health* 2010;13(5) doi: https://doi.org/10.1111/j.1524-4733.2010.00700.x
46. Latimer N. NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. NICE Decision Support Unit. 2013. *NICE DSU technical support document* 2013
47. Ruiz-Morales JM. First-line sunitinib versus pazopanib in metastatic renal cell carcinoma: Results from the International Metastatic Renal Cell Carcinoma Database Consortium. *European Journal of Cancer* 2016;65:102-08. doi: http://dx.doi.org/10.1016/j.ejca.2016.06.016
48. National Institute for Health and Care Excellence. Single technology appraisal (STA). Pazopanib (Votrient®) for the first-line treatment of patients with advanced renal cell carcinoma (RCC). Manufacturer submission. 16 April 2010 04/16/2010. https://www.nice.org.uk/guidance/ta215/documents/renal-cell-carcinoma-first-line-metastatic-pazopanib-manufacturer-submission-submission2 (accessed January 2018).
49. National Institute for Health and Care Excellence. TA169. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma 25 March 2009 [Available from: https://www.nice.org.uk/guidance/ta169 accessed January 2018.
50. Amdahl J. Cost-effectiveness of pazopanib compared with sunitinib in metastatic renal cell carcinoma in Canada. *Current Oncology* 2016;23(4):e340-e54.
51. Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med* 2015;373(19):1814-23. doi: 10.1056/NEJMoa1510016 [published Online First: 2015/09/26]

52. Curtis L, Burns A. Unit Costs of Health and Social Care 2017. 2017 ed. Canterbury, Kent: Personal Social Services Research Unit. University of Kent, 2017.
53. NHS Improvement. NHS Reference Costs 2016/2017. 2017 [Available from: https://improvement.nhs.uk/resources/reference-costs/ accessed 10/4/2018.
54. National Institute for Health and Care Excellence. Everolimus for advanced renal cell carcinoma after previous treatment. Technology appraisal guidance (TA432). London: NICE, 2017.
55. National Institute for Health and Care Excellence. Lenvatinib with everolimus for previously treated advanced renal cell carcinoma Technology appraisal guidance (TA498). London: NICE, 2018.
56. National Institute for Health and Care Excellence. Axitinib for treating advanced renal cell carcinoma after failure of prior systemic treatment. Technology appraisal guidance (TA333). London: NICE, 2015.
57. National Institute for Health and Care Excellence. Nivolumab for previously treated advanced renal cell carcinoma. Technology appraisal guidance (TA417). London: NICE, 2016.
58. Hudes G, Carducci M, Tomczak P, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. *New England Journal of Medicine* 2007;356(22):2271-81. doi: DOI: 10.1056/NEJMoa066838
59. Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. *The Lancet Oncology* 2016;17(7):917-27. doi: https://doi.org/10.1016/S1470-2045(16)30107-3
60. Rini BI, Halabi S, Rosenberg JE, et al. Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206. *Clinical Oncology* 2008;26(33):5422. doi: https://dx.doi.org/10.1200%2FJCO.2008.16.9847
61. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. London: Nuffield Trust, 2014.
62. Inflation.EU. Average inflation rate in Great Britain [Available from: http://www.inflation.eu/inflation-rates/great-britain/historic-inflation/cpi-inflation-great-britain.aspx. accessed 9/8/16.
63. Rini. BI, Vogelzang. NJ. Future Challenges for Drug Development in Renal Cell Carcinoma. *Journal of Clinical Oncology* 2017;35(6):577-79. doi: 10.1200/jco.2016.71.0673

64. Motzer RJ, Hutson TE, Tomczak P, et al. Sunitinib versus Interferon Alfa in Metastatic Renal-Cell Carcinoma. *New England Journal of Medicine* 2007;356(2):115-24. doi: 10.1056/NEJMoa065044
65. Ko JJ, Choueiri TK, Rini BI, et al. First-, second-, third-line therapy for mRCC: benchmarks for trial design from the IMDC. *British Journal Of Cancer* 2014;110:1917. doi: 10.1038/bjc.2014.25
66. Choueiri TK, Halabi S, Morris MJ, et al. Reply to B. Rini et al and S. Buti et al. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2017;35(16):1859-60. doi: 10.1200/jco.2017.72.2629 [published Online First: 2017/05/27]
67. Schnadig ID, Hutson TE, Chung H, et al. Dosing patterns, toxicity, and outcomes in patients treated with first-line sunitinib for advanced renal cell carcinoma in community-based practices. *Clinical genitourinary cancer* 2014;12(6):413-21. doi: 10.1016/j.clgc.2014.06.015 [published Online First: 2014/08/29]

9 APPENDICES

9.1 ERG critical appraisal of the ITC

Criterion	ERG assessment
ITC purpose	
1. Are the ITC results used to support the evidence for the clinical effectiveness of the intervention?	Yes, for the comparison of cabozantinib with pazopanib.
2. Are the ITC results used to support the evidence for the cost-effectiveness of the intervention?	Yes. OS and PFS results from the ITC are used directly in the economic model to inform the estimates of cost effectiveness.
Evidence selection	
3. Are inclusion/exclusion criteria adequately reported?	Yes. CS Table 4 lists the inclusion criteria. These criteria include a broader list of treatments than in the NICE scope. The CS notes that the systematic review was conducted from a global perspective and consequently included additional comparator treatments not specified in the NICE scope. Subsequent restriction to only comparator treatments in the scope resulted in inclusion of 2 studies (n=9 records), the CABOSUN trial and the COMPARZ trial.
4. Is quality of the included studies assessed?	Yes, for the 2 studies in the restricted ITC network (Table 15 in Appendix D1.3, and Figure 41 and 42 in Appendix D1.1), using the standard criteria recommended by NICE.
Methods – statistical model	
5. Is the statistical model described?	Yes. Three types of statistical method are used: (1) Indirect comparison of parametric survival curves using methodology developed by Ouwens et al (2010) (2) Parametric models with fractional polynomial distributions using methodology developed by Jansen et al (2011). (3) A network meta-analysis supplementary method comparing hazard ratios using a fixed effects model, for intermediate risk and poor risk subgroups. Methods 1 and 2 are used to inform the economic model.
6. Has the choice of outcome measure used in the analysis been justified?	Yes, OS and PFS are key outcomes in cancer survival modelling.
7. Has a structure of the network been provided?	Diagrams illustrating the networks are provided in the CS: Figure 9 shows the primary evidence network for potential meta analysis (i.e. based on the broader inclusion criteria). CS Figure 11 shows the restricted evidence network containing the 2 included RCTs. CS Appendix D1.1 shows the networks used in the NMA supplementary method of HRs (CS Figures 43 to 48).
8. Is homogeneity considered?	Yes, discussed in CS section B.2.9 and Appendix D1.1. A feasibility assessment is described to assess differences in study and patient characteristics within and between treatment comparisons. CS Table 22 tabulates risk category and performance status details between the 2 included trials.

9. Are the studies homogenous in terms of patient characteristics and study design?	No. The CABOSUN trial included only patients at intermediate or poor risk, whilst the COMPARZ study included patients with favourable, intermediate and poor risk classifications. The distribution of patients between risk classifications is different between the two trials. The CS acknowledges that the differences in distribution of risk category is the variable that most affects survival. There were slight differences between trials in the number of metastatic sites detected (≥3 sites: 32% to 41% by treatment arm in CABOSUN; 42% to 44% by treatment arm in COMPARZ). (CS Appendix Table 11). Just over a third of patients in CABOSUN had bone metastases at baseline (36%-37% by trial arm) compared to 15%-20% (by trial arm) of patients in COMPARZ. CABOSUN was a small phase II RCT (n=157 patients), whilst COMPARZ was a larger phase III RCT (n=1110 patients randomised).
10. If the homogeneity assumption is not satisfied, is clinical or methodological homogeneity across trials in each set involved in the indirect comparison investigated by an adequate method? (e.g. sub group analysis, sensitivity analysis, meta-regression)	The network meta-analysis supplementary method presented in Appendix D conducted separate NMAs in intermediate risk and poor risk subgroups, including comparators outside of the scope. This NMA is reported to be additional to the Ouwens et al and fractional polynomial analyses, specifically to explore the results of subgroup analyses compared to the overall study populations. However, unlike the other two analyses, this method does assume proportional hazards.
11. Is the assumption of similarity stated?	Yes – the CS discusses similarity ('assessment of heterogeneity' in Section B.2.9), describing the similarities and differences between the two trials, but does not explicitly state whether the assumption of similarity holds. In Appendix D1.1 it is stated that the "populations in CABOSUN and COMPARZ are different".
12. Is any of the programming code used in the statistical programme provided (for potential verification)?	Yes, provided in CS Appendix D1.1.
Sensitivity analysis	
13. Does the study report sensitivity analyses?	No.
Results	
14. Are the results of the ITC presented?	Yes. Most of the detail is in Appendix D1.1, with presentation of a series of graphs showing fitted survival curves for the Ouwens et al model and the fractional polynomials models, for both OS and PFS and for random and fixed effects models, where conducted (Figures 1 to 40). However, the CS did not report hazard plots depicting the time-varying hazard ratios and their credible intervals from the fractional polynomials models. The ERG requested these plots from the fractional polynomials analysis from the company (clarification question A22 and A23).

15. Does the study describe an assessment of the model fit?	Yes. Fit statistics for the Ouwens et al and the fractional polynomial methods, for OS and PFS, are presented in CS section B.2.9 (Tables 23 and 24). The Deviance Information Criteria (DIC) was used to select the model with the best fit, with a lower posterior mean DIC indicating a better fit. This is a standard approach to assessing Bayesian model fit. The CS does not report any other considerations in relation to model fit (e.g. plausibility of modelled distribution).
16. Has there been any discussion around the model uncertainty?	No. The ERG requested the company to provide the credible intervals for the time-varying hazard ratios estimated by the fractional polynomial model (clarification question A22) to assess the degree of uncertainty. The Ouwens et al models were conducted using fixed effect and random effects, and the fractional polynomials models were conducted using only fixed effects. The company were requested to supply the random effects fractional polynomial model (clarification question A23).
17. Are the point estimates of the relative treatment effects accompanied by some measure of variance such as confidence intervals?	No. However, the Ouwens et al and fractional polynomial methods do not estimate a single point estimate, such as a constant hazard ratio. For example, the fractional polynomial method estimates time-varying hazards over time. The ERG requested the company to provide hazard ratios and credible intervals for each interval of the follow-up time period for the fractional polynomial models (clarification question A22).
Discussion - overall results	
18. Does the study discuss both conceptual and statistical heterogeneity?	Yes. Conceptual (clinical) heterogeneity is discussed (see above). Statistical heterogeneity was not relevant as the ITC in the restricted network included only two trials, linked together by a common comparator arm.
Discussion - validity	
19. Are the results from the indirect/NMA compared, where possible, to those just using direct evidence?	No. This was not necessary as there are no comparisons informed by both direct and indirect evidence.

9.2 Critical appraisal of the COMPARZ trial

The table below presents the company's and the ERG's critical appraisal of the COMPARZ trial.[34]

NICE QA Criteria for RCT	CS response	ERG response
1. Was the method used to generate random allocations adequate?	Yes	Unclear
The trial publication states that patients were randomly assigned to one of the two study drugs in a 1:1 ratio in permuted blocks of four. The CSR states that the randomisation schedule was generated by GSK Statistics and Programming Department (page 45). However, it does not state the exact method used to generate the schedule.		
2. Was the allocation adequately concealed?	Not clear	Yes

Comments: The CSR states that an interactive voice response system was used (section 5.3). All patients were entered into this system after baseline assessment and the randomisation schedule was then generated centrally. It appears that study sites called the interactive voice system to request randomisation when required. Thus sites could not have known in advance the next random allocation in the sequence.		
3. Were the groups similar at the outset of the study in terms of prognostic factors, e.g. severity of disease?	Yes	Yes
Comments: There do not appear to be any notable differences between the groups in demographic or clinical characteristics (Supplementary Table S3 to the trial journal publication ([34]).		
4. 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	No
Comments: The trial was open-label. However, imaging data were re-evaluated by an independent review committee whose members were unaware of the treatment assignments to assess the primary end point and tumour response.		
5. Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No	No
Comments: The number of treatment discontinuations was similar between the two groups, and the reasons for discontinuations were broadly similar (Supplementary Figure S2 to the trial journal publication[34])		
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
Comments: The ERG checked the objectives (outcomes) stated in the CSR and outcome data are reported for each of them.		
7. 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 Unclear
Comments: Efficacy data were analysed in the intention-to-treat population (all patients who underwent randomisation). However, the ERG notes that for patient reported outcomes (HRQoL and symptoms) the number of patients analysed is lower than the number randomised. It is not clear how missing data were handled (see Table 2 in the trial journal publication.[34]		

9.3 Description and critique of ITC method 3: Network meta-analysis supplementary method

CS Appendix D1.1 reports brief details of what the CS describes as a supplementary NMA of cabozantinib compared to sunitinib, pazopanib, interferon-alfa, sorafenib, bevacizumab in combination with interferon-alfa, temsirolimus, and tivozanib. The CS reports that non-scope treatments were included in this network because the analysis was conducted for a non-English perspective.

The CS states that Kaplan-Meier data results were not available for intermediate and poor prognostic risk groups separately from the ITT population (CS Section B.2.9, page 56). To further explore the impact of differences in the subgroup data, an additional analysis was carried

out on hazard ratios (HRs). Unlike Kaplan-Meier data, HRs were available by subgroup and these were compared despite the violation of the proportional hazards assumption (see Appendix D for further details).

Separate evidence networks were constructed based on RCC risk groups: intermediate risk, poor risk, and the overall population of patients. The networks include comparators inside and outside of the NICE appraisal scope, and they vary in size according to population group (e.g. overall population, risk subgroup) and outcome measure. The CS describes the set of studies as heterogeneous in terms of RCC risk groups, with some studies including patients with favourable RCC risk. The CS does not provide any further details of the characteristics of the included studies, but does tabulate the OS and PFS HRs for the ITT populations, and intermediate and poor risk subgroups for each of the trials (CS Appendix D1.1 Table 14). The results of the NMA are presented as a series of fixed effect forest plots showing the HRs for each of the treatments compared to cabozantinib (CS Appendix D1.1 Figure 49 to Figure 51).

The ERG urges caution in the interpretation of these analysis as they assume that the proportional hazards assumption holds in the CABOSUN trial (and other trials in the network), yet as discussed above, this assumption is not supported by the OS curves in the trial. Furthermore, no assessment of heterogeneity or consistency has been provided for the trials in the networks, and the networks use data from subgroups of the randomised patient populations (the size of which are unspecified in the CS), therefore can be considered observational evidence, and likely underpowered due to small sample sizes. Furthermore, very little information is given on the statistical methods used to conduct this analysis.

9.4 Additional results of the ITC

Fixed effect ITC fractional polynomial model – additional results

In section 3.3.7 of this report we reported the results of the best fitting fractional polynomial models. Here we summarise the results of the other fractional polynomial models tested. There were some differences in results between the different fractional polynomial models:

Progression free survival

- **First order model results**. In three of the models (P=0; P=0.5; P=1) there was a slight decline in the time-varying HR curves over time, from around 0.5 to around 0.3-0.4 (NB.

The ERG was unable to cross-check the HR plots with the tabulated HRs for first order PFS pazopanib as Table 5 appears to be a duplicate of Table 3, which is the tabulated HRs for first order OS pazopanib (clarification question A22)). Credible intervals tended to increase markedly over time and exceeded 1. In the other two models (P=-1; P=-0.5), the HR curves were flat at around 0.5 for the entire follow-up period, indicating little change in HRs over time.

- **Second order model results**. In all models the time-varying HRs for both comparisons increase sharply from zero within the first three months to reach a plateau of around 0.5, then decline slightly over time to around 0.4. Credible intervals tended to increase markedly over time in all models and exceeded 1, though the intervals in the best fitting model (p1=-1, p2=-1) are less wide than the other models. The results of the best fitting fractional polynomial as used in the economic model are therefore consistent with the other second order models, though with less uncertainty.

Overall survival

- **First order model results**. In most of the first order fractional polynomial models the time-varying HRs curves are reasonably straight over time (at around 0.7-0.8), indicating a constant HR. The exception is first order fractional polynomial model p=1 in which the curves decline slightly over time from around 0.8-1.0 to around 0.6. Fractional polynomial first order model (p=-0.5) appears to be an outlier as pazopanib has a slightly higher HR compared to sunitinib by an order of approximately 0.1 (around 0.6 compared with around 0.5, respectively) and the credible intervals are wider than all the other first order models.
- **Second order model results.** The second order fractional polynomial model curves have a distinctly different shape to the first order curves. As was the case for PFS, the time-varying HRs for both comparisons increase sharply from zero within the first six months to reach a plateau, then decline slightly over time. The exception is the best fitting fractional polynomial model (p1=-1, p2=-1) where the HRs remain generally constant (and higher than the other models) once they have peaked.

Random effects fractional polynomial model results

The fractional polynomial ITC results presented in the CS were based on a fixed effect model. For comparison, the company were asked to provide fractional polynomial results based on a random effects model (clarification question A23). The ERG crosschecked the results of the fixed effect and random effects fitted fractional polynomial curves (for the restricted network

only). In all but one of the fractional polynomial models, the results appeared similar between the fixed effect and the random effects models. The exception was the OS 1st order (p=-0.5) model where the random effects model (Figure 39, clarification question A23) had higher curves for all three treatments compared to the fixed effect model (CS Appendix D1.1 Figure 24). It is not clear why this is the case. Importantly, the fitted curves for random effects and fixed effect models in the best-fitting fractional polynomial model (used to inform the economic model) were similar to each other, indicating that the inclusion of additional evidence did not change the results.

ITC results for the wider evidence network

The company were asked to provide ITC results based on the wider network of 13 RCTs that included studies of additional treatments not within the scope of the appraisal (clarification question A26) (see section 3.1.7.1 for a discussion of this network). The aim was to check whether the results for the comparison between cabozantinib, sunitinib and pazopanib were different when a wider network containing other treatment comparisons was used.

The company point out that there is considerable clinical heterogeneity in this network, citing the TARGET study of sorafenib versus placebo as comprising a mostly pre-treated population. They also mention that there were differences in the extent of patient crossover in some trials. The company has not presented tabulated characteristics of these studies to allow an assessment of clinical heterogeneity, but the ERG agrees that it is plausible that clinical heterogeneity would exist in this wider set of studies.

The ERG cross-checked the results of the wider and the restricted networks for the Ouwens et al ITC fixed effect and random effects models. The results were similar in all cases except for the exponential model where there were bigger differences in the fitted survival curves between pazopanib and sunitinib (whereas in the restricted network they were similar). The reason for this disparity between the networks is not clear. Results from Gompertz survival models in the wider network were not supplied in response to clarification question A26 so the ERG are unable to check the consistency of results for this model between the networks.

The ERG cross-checked the results of the wider and the restricted networks for the fixed effects fractional polynomials models. In all but one of the models, the results appeared similar between the wider and restricted networks. The shape of the fitted PFS survival curves for

cabozantinib, sunitinib and pazopanib in first order model (p=-0.5) of the restricted network (CS Appendix D1.1 Figure 34) did not correspond to the corresponding curves in the wider network (Figure 131, clarification question A26). It is not clear why this is the case. Also, one of the fractional polynomial second order models (P1=-0.5, P2=0) based on the wider network did not converge, so it is not possible to compare its results with the corresponding model in the restricted network. Importantly, the fitted curves for the wider and the restricted networks in the best-fitting fractional polynomial model (used to inform the economic model) were similar to each other, indicating that the inclusion of additional evidence did not change the results.

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Cabozantinib for untreated metastatic renal cell carcinoma [ID1208]

You are asked to check the ERG report from Southampton Health Technology Assessments Centre to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on 20 April** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

In this pro-forma, you will find details of three factual inaccuracies, three points of clarifications, two CIC unmarking and nine typos.

Issue 1 Factual inaccuracy - OS dataset used in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 21, Weaknesses and areas of uncertainty Correction of the following ERG statements which are factually incorrect: *'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 does not explicitly state which OS dataset was used to inform in the model, but the January 2017 KM plot is reproduced in the CS economic chapter and KM data provided by the company in response to a clarification question also relates to this earlier cut-off). This suggests that the model may over-estimate the survival advantage for cabozantinib over sunitinib'*.	Please amend both statements to clarify that the CS did state which OS dataset was used to inform the model. Proposed text is as follows: *'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 ~~does not explicitly~~ states ~~which~~ that the OS January 2017 dataset was used to inform in the model~~, but the January 2017~~ and the KM plot is reproduced in the CS economic chapter and KM data provided by the company in response to a clarification question ~~also relates to this earlier cut-off).~~ This suggests that the model may over-estimate the survival advantage for cabozantinib over sunitinib'*.	Factually inaccurate statement. Throughout the CS it is explicitly stated that the OS results contained within the CSR were used to inform the economic model, and specific data is provided in the clinical section (see below): CS Page 39, Overall survival (data cut off 13 January 2017) states '*OS data are summarised in Table 13 and Figure 6 (Note: these data are used to inform the economic model.)*'	Corrected

Issue 2 Factual inaccuracy - OS dataset used in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 82, Section 4.3.4.1 Overall Survival Correction of the following ERG statement which is factually incorrect: *'The CS does not state which dataset was used, but the January 2017 KM plot is reproduced in the economic chapter (CS B.3.3 Figure 13) and KM data provided by the company in response to a clarification question also relates to this earlier cut-off.*'	Please amend the statement to clarify that the CS did state which OS dataset was used to inform the model. Proposed text is as follows: *'The CS ~~does not~~ stated that the data cut off January 2017 ~~which~~ dataset was used (CS page 39, Table 13 and Figure 6)~~, but the January 2017~~. The KM plot ~~is~~ 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~~s~~ to this ~~earlier~~ cut-off.*	Factually inaccurate statement. Throughout the CS it is explicitly stated that the OS results contained within the CSR were used to inform the economic model, and specific data is provided in the clinical section (see above response to Issue 1).	Corrected as suggested

Issue 3 Factual inaccuracy

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 57, Section 3.1.7.6 Choice of fractional polynomial model Factual inaccuracy The following statement is factually incorrect: *'The CS does not state whether any other considerations were taken into account in the choice of model, such as clinical plausibility*	The following amendment, which reflects the text in the CS submission, is proposed: *'~~The CS does not state whether any other considerations were taken into account in the choice of model, such as clinical plausibility with respect to the OS and PFS estimates generated~~ The CS states that in order to select the best survival model fit, the algorithm (SMEEP) as described in the NICE DSU Technical Support Document 14 was followed.*	Factual inaccuracy which needs amending to accurately reflect the CS	It is our understanding that the plausibility of different extrapolations with oncologists was done for the survival models generated by the Ouwens et al method. No mention is made regarding plausibility of model fit for the fractional polynomial method. We have not made any changes to this text.

with respect to the OS and PFS estimates generated.' Information is provided in the CS submission Section B3.3, Clinical parameters and variables.	*This included the use of Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC) statistics visual inspection of the curves. The plausibility of different extrapolations was also assessed by visual inspection by oncologists currently practising within the NHS in England. The most appropriate model was selected based on a combination of all these factors.'*		

Issue 4 Points of clarification

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 22, Summary of additional work undertaken by the ERG, Bullet point 'Method of fitting OS curves' Page 111, Table 26 Column 'Reason for analysis', Row OS Curves, Simple indirect comparison Points of clarification The following ERG statements would benefit from clarification, in order to highlight that, based on the evidence from COMPARZ, which demonstrated that pazopanib and sunitinib were the same, the CABOSUN data was used for pazopanib (making it equivalent to sunitinib). Page 22: *Finally, we assume*	In order to accurately reflect and avoid any misinterpretation, the following amendments are proposed: Page 22 *'Finally, based on the relationship shown in COMPARZ, we assume equivalent OS for pazopanib and sunitinib, ~~based on~~ using the ~~results~~ evidence ~~of~~ from ~~COMPARZ~~ CABOSUN'* Page 111 *'OS assumed equal for pazopanib and sunitinib, based on the relationship in COMPARZ'*	To avoid misinterpretation and to accurately clarify the OS assumption.	Amended as suggested

equivalent OS for pazopanib and sunitinib, based on the results of COMPARZ' Page 111 '*OS assumed equal for pazopanib and sunitinib, based on COMPARZ'*			

Issue 5 Point of clarification

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 55, Section 3.1.7.4 ITC: comparison of parametric survival curves Point of clarification The following sentence would benefit from additional text to add further clarification. '*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 deaths for each interval and patients at risk at the beginning of the interval*'	The following amendment is proposed: '*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 <u>progression events or</u> deaths for each interval and patients at risk at the beginning of the interval*'	Point of clarification	Amended as suggested

Issue 6 Point of clarification

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Economic model 'ID1208_Cabo - ERG analysis (CIC) - SA 10042018 (CiC)': Point of clarification The ERG proposed that, due to the differences in the risk groups between the two trials used in the ITC, pazopanib should be assumed to have equal efficacy to sunitinib (from relationship shown in COMPARZ). Thus, OS and PFS data from CABOSUN are used for pazopanib. However, the subsequent treatments for pazopanib have not been changed, thus assuming that pazopanib and sunitinib patients will have the same life expectancy, but with different subsequent treatments.	We acknowledge that the ERG has proposed different subsequent treatment data for the scenario analyses, but we believe that for the base case analysis, subsequent treatment lines following pazopanib should be the same as those following sunitinib.	To align the assumption on second line treatments, with that made on equal efficacy of first line treatments.	Noted, but not a factual error. No change made.

Issue 7 Unmarking of CIC data

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 9, Summary of submitted clinical effectiveness evidence Unmarking of information	The CIC mark up of the following text can be removed: *'Patient cross-over between trial arms was not permitted during the trial, however, upon*	Removal of CIC marking as data is not CIC	Corrected

incorrectly marked as CIC	*disease progression patients in both arms received subsequent systemic non-radiation anti-cancer treatments (cabozantinib group 57%; sunitinib group 58%)*'.		

Issue 8 Unmarking of CIC data

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 123, Section 5, Table 35 Unmarking of CIC data	The CIC mark up of the following text can be removed: *'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'.*	Data is no longer CIC	Amended

Issue 9 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 8, Scope of the company submission Typographical error	The sentence *'The recommended dose is 60 mg once daily, with lower dose adjustments recommend to manage adverse reactions'*.	Typographical error	Corrected

	needs to be corrected to *'The recommended dose is 60 mg once daily, with lower dose adjustments recommended to manage adverse reactions'*.		

Issue 10 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 12, Indirect treatment comparison Typographical error	Sentence '*Results for OS from the best-fitting PFS fractional polynomial model (which informs the economic model base case) show'* needs to be corrected to *'Results for OS from the best-fitting ~~PFS~~ fractional polynomial model (which informs the economic model base case) show:*'	Typographical error	Corrected

Issue 11 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 21 Summary of additional work undertaken by the ERG; Page 114, text above table and Table 28 Cost effectiveness: Company base-case analyses (ERG Corrected); Page 118 Table 30 Scenario analysis: Company ITC base case (ERG corrected), vs. sunitinib; Page 126 7.2 Summary of cost	The figure *'£26,182'* is incorrect and needs to be amended to *'£26,185'*	Typographical error	The correct figure is actually £26,182 for our simple corrected version of the company model, without any additional ERG analysis. The figure of £26,185 comes from our version of the analysis, set up to run our preferred analysis and scenarios. The results are very

effectiveness issues Typographical error			similar, but not exactly the same when our model is set up to run the company's base case. The reason for this is that our version of the model includes lenvatinib and everolimus instead of interferon as a subsequent treatment option, so that we could include this in our scenario analysis on the subsequent treatment utilisation. This makes very little difference for the company's base case, because only assume 1% of patients had interferon after failure of cabozantinib in the trial, and none after sunitinib or pazopanib.

Issue 12 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Pages 26, Section 2.1.2.1 Staging and prognosis Typographical error	*'Poor - > 3 factors'* needs to be corrected to *'Poor - ≥ 3 factors'*	Typographical error to be accurate, this was incorrect in the CS	Corrected

Issue 13 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Pages 41 and 42, Section 3.1.5	Page 41, second bullet point text '*For the*	Typographical error	Corrected

Description and critique of company's outcome selection Typographical error	*regularity submission'* needs to be corrected to '*For the regulatory submission'* Final sentence on page 42 '*HRQoL in cancer trials it is an important outcome...*' needs to be corrected to '*HRQoL in cancer trials ~~it~~ is an important outcome...*'		

Issue 14 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 51, Clinical heterogeneity Typographical error	Text *'(NB The data for Karnofsky performance status 70 to 80 and 80 to 100 are the wrong way round in CS Table 11)'* needs to be amended to *'(NB The data for Karnofsky performance status 70 to 80 and 80 to 100 are the wrong way round in CS Appendix Table 11)'*	Typographical error and in order to refer to the correct table	Corrected

Issue 15 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 55, Section 3.1.7.4 ITC: Comparison of parametric survival curves Typographical error	First sentence '*conducing*' needs to be amended to '*conducting*'	Typographical error	Corrected

Issue 16 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 111, Table 26 Column	'*HR=80'* is incorrect and needs to be amended	Typographical error	Corrected

'Reason for analysis', Row OS Curves, Simple indirect comparison Typographical error	to *'HR=0.80'*		

Issue 17 Typographical error in economic model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Economic model 'ID1208_Cabo - ERG analysis (CIC) - SA 10042018 (CiC)' Sheet: Variables Cell: L45	Cell L45 in 'Variables' sheet should read: '=IF(PSA_Active,AC45,IF(AND(Tornado_Active,AK45<>""),AK45,F45))'	We thank the ERG for their corrections to the model. The proposed amendment is intended to build on these and highlight a further amendment to cell L45 to ensure appropriate implementation.	Acknowledged. This cell was set up correctly in our corrected version of the company model, but we forgot to change it in the ERG preferred analysis version of the model. This doesn't make any difference to any results in the ERG report.

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

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors 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 Chorozoglou, Senior Research Fellow (Health Economics)

Correspondence to Dr Jonathan Shepherd
Southampton Health Technology Assessments Centre (SHTAC)
Wessex Institute
Alpha House
Enterprise Road, University of Southampton Science Park
Southampton SO16 7NS
www.southampton.ac.uk/shtac

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 (≥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 ≥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

The CS reports the results of the ITC as fitted survival curves for the outcomes of OS and PFS for all three treatments (cabozantinib, sunitinib, pazopanib), based on fixed effect and on random effects, for each of the five parametric distributions generated by the Ouwens et al method. For each of the analyses cabozantinib had a higher survival estimate than sunitinib or pazopanib. The sunitinib and pazopanib curves were similar to each other in shape and position, indicating similar effectiveness between these two treatments.

The CS presents fitted fractional polynomial survival curves for the outcomes of OS and PFS for all three treatments, based on fixed effects for first and second order models. On request the company also supplied HR plots with credible intervals for each fractional polynomial model to allow visual inspection of the time-varying HR curves. Results for PFS from the best-fitting fractional polynomial model (which informs the economic model base case) show:

- The HR for pazopanib peaks at month four ([redacted]) and declines slightly during the rest of the follow-up period. The HR for sunitinib peaks at month six ([redacted]) and declines slightly during the remainder of the follow-up period.
- The credible intervals increase over the follow-up period, with the upper bound increasing to include 1 after month 19 for pazopanib, and after month 11 for sunitinib.
- The time-varying PFS HRs for cabozantinib versus sunitinib generated by this fractional polynomial model compare broadly with the constant HR reported in the CABOSUN trial (0.48 (95% CI 0.31, 0.74)), though with greater uncertainty (wide credible intervals).

Results for OS from the best-fitting fractional polynomial model (which informs the economic model base case) show:

- The HR for pazopanib starts to peak at month nine, and declines slightly after month 19 ([redacted]). The HR for sunitinib begins to plateau at month 13 and peaks at month 30 where it remains for the rest of the follow-up period ([redacted]).
- The credible intervals widen during the course of the follow-up period, and include 1 at all time points.
- The time-varying OS HRs for cabozantinib versus sunitinib generated by this fractional polynomial model compare broadly with the constant OS HR reported in the CABOSUN trial (0.80 (95% CI 0.53, 1.21), though with greater uncertainty (wide credible intervals).

Across the other fractional polynomial models (first and second order), the time-varying HR curves for cabozantinib versus sunitinib and cabozantinib versus pazopanib have a similar

- 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 - ≥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).[7] 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.[2]

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.[9] 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.[10]

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.[10 11] As might be expected, evidence shows that the effects of disease progression in these patients is linked to a deterioration in HRQoL.[12 13 14 15]

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.[16] 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 ≥ 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.[33] 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;[24]
- 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).[24] Results of the updated analysis are also reported in the CSR and in a conference presentation.[30]

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 EGOG 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%.[37] 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.[36] 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[36] 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[36] 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.[36]

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.[47] 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

	Preferred assumptions	Scenarios	Reason for analysis
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 >=5% incidence	Range of disutilites, 8 week duration and >=2%	Again, we follow the company approach, but conduct additional analyses to test the sensitivity of the model to adverse events.

	Preferred assumptions	**Scenarios**	**Reason for analysis**
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	**Direct**: Exponential (Weibull & Gompertz) **ITC**: FP model with P1=P2=-1 (exponential; Weibull; Gompertz; and FP P1=-0.5, P2=0 & P1=-1, P2=0)	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. Given these reservations, the exponential, Weibull and Gompertz are reasonable for the direct analysis. For the ITC, the exponential and FP P1=P2=-1 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.	Simple indirect comparison assuming: • Sunitinib OS curve based on company's exponential fit to CABOSUN; • Cabozantinib calculated from sunitinib curve and HR from July 2017 CABOSUN results; • Pazopanib curve assumed equal to sunitinib (based on COMPARZ results).
PFS curves	**Direct**: lognormal (Exponential, Weibull & Gompertz) **ITC**: FP P1=P2=-1 (exponential, Weibull and Gompertz)	CABOSUN PFS analysis is more mature. ITC is subject to uncertainty due to differences in trial populations, unclear if similarity assumption is met. 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.	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.
TTD curves	**Direct**: lognormal (exponential, Weibull, Gompertz & gamma).	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.	Lognormal for base case, and all other distributions in scenario analysis.

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)

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	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.
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	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 ([redacted] 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