NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Contents:

The following documents are made available to stakeholders:

Access the **final scope and final stakeholder list** on the NICE website.

1. **Company submission** from Ipsen:
   a. Full submission
   b. Summary of Information for Patients (SIP)

2. **Clarification questions and company responses**
   a. Main response
   b. Clarification follow up responses
   c. Clarification response scenario results

3. **External Assessment Report** prepared by School of Health and Related Research (ScHARR)

4. **External Assessment Report – factual accuracy check**

Post-technical engagement documents

5. **Technical engagement response from company**

6. **Technical engagement responses and statements from experts:**
   a. Dr Sath Nag – clinical expert, nominated by Society for Endocrinology
   b. Professor Jonathan Wadsley – clinical expert, nominated by NCRI-ACP-RCP-RCR

7. **Technical engagement responses from stakeholders:**
   a. NCRI-ACP-RCP-RCR

8. **External Assessment Report critique of company response to technical engagement** prepared by School of Health and Related Research (ScHARR)

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Document B

Company evidence submission

September 2022

File name	Version	Contains confidential information	Date
ID4046 Cabozantinib DTC Company Evidence Submission Doc B Final V1.0 16Sept2022 AIC_CIC	**V1.0**	**Yes**	**16th September, 2022**

Contents

Tables and figures

Abbreviations

AE	Adverse Event
ALT	Alanine aminotransferase
AS	European age-standardised
ATA	Adequate Tumour Assessment
BIRC	Blinded Independent Radiology Committee
BOR	Best Overall Response Rate
BSC	Best Supportive Care
CCO1	Clinical Cut-Off 1
CCO2	Clinical Cut-Off 2
CI	Confidence Interval
CR	Complete Response Rate
DOR	Duration of Objective Response
DTC	Differentiated Thyroid Carcinoma
ECG	Electrocardiogram
EMA	European Medicines Agency
ESMO	European Society for Medical Oncology
ETA	European Thyroid Association
HRQoL	Health Related Quality of Life
IF	Information Fraction
KM	Kaplan-Meier
MET	Mesenchymal Epithelial Transition Factor Receptor
NCCN	National Comprehensive Cancer Network
NICE	National Institute of Health and Care Excellence
NMA	Network meta-analysis
NPACT	Non-Protocol Anticancer Therapy
NTRK	Neurotrophic Tyrosine Receptor Kinase
ORR	Overall Response Rate
OS	Overall Survival
PD	Progressed Disease
PFS	Progression-free Survival
PR	Partial Response Rate
RAI	Radioactive Iodine
RECIST	Response Evaluation Criteria in Solid Tumours
RPSFT	Rank-preserving Structural Failure Time
RTKs	Receptor Tyrosine Kinases
SoD	Sum of target lesion Diameters
TKI	Tyrosine Kinase Inhibitor
TSH	Thyroid-Stimulating Hormone
VEGFR	Vascular Endothelial Growth Factor Receptor

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

B.1.1 Decision problem

B.1.1.1 Population

The marketing authorisation is: “CABOMETYX is indicated as monotherapy for the treatment of adult patients with locally advanced or metastatic differentiated thyroid carcinoma (DTC), refractory or not eligible to radioactive iodine (RAI) who have progressed during or after prior systemic therapy.”[1]

The population defined in the final scope is adults with locally advanced or metastatic differentiated thyroid carcinoma, whose disease is refractory to, or who are unsuitable for radioactive iodine, and whose disease has progressed during or after prior systemic therapy. This submission covers the technology’s full marketing authorisation for this indication.

The decision problem addressed within this submission is consistent with the NICE final scope for this appraisal as outlined in Table 1.

B.1.1.2 Comparator

In the COSMIC-311 trial the control arm was matched placebo, as at the time of clinical trial design there were no other treatments indicated for RAI-refractory DTC. According to the European Society for Medical Oncology (ESMO) treatment guidelines, lenvatinib or sorafenib should be considered as the standard first-line treatment systemic therapies for RAI-refractory DTC.[2] For advanced/metastatic RAI-refractory DTC, ESMO suggests cabozantinib and lenvatinib as two potential choices for second-line treatment of patients who have progressed on sorafenib. However, the sequencing pathway cannot be confirmed due to the current available evidence, with ESMO stating that “the decision should be individualised for each patient considering the likelihood of response and safety profile of the drug”. Therefore, the optimal sequencing pathway will be unique to each DTC patient.[3]

NICE technology appraisal 535 (TA535)[4] recommends lenvatinib and sorafenib, which inhibit multiple receptor tyrosine kinases including vascular endothelial growth factor receptors (VEGFR) only for people who have not had tyrosine kinase inhibitors before, or who must stop them early because of tolerability (specifically, toxicity that cannot be managed by dose delay or dose modification). This is because there is not enough clinical evidence and no cost-effectiveness evidence to determine whether the treatments are effective when used sequentially.

National Health Service (NHS) England Cancer Drugs Fund (CDF) criteria for use state: “Sequential use of lenvatinib and then sorafenib is only funded if the patient has to discontinue lenvatinib because of intolerance within 3 months of its start and if the disease has not progressed whilst the patient is on lenvatinib. The use of lenvatinib after disease progression on or after sorafenib is not funded and vice versa.”[5]

The only second-line treatment that has recently been recommended by NICE which could be used to treat RAI-refractory DTC is selpercatinib (TA742).[6] It is recommended for use within the CDF, as an option for treating advanced rearranged during transfection (RET) fusion-positive thyroid cancer in adults who need systemic therapy after sorafenib or lenvatinib. Additionally, patients with a neurotrophic tyrosine receptor kinase (NTRK) fusion could also be retreated with entrectinib (TA644)[7] or larotrectinib (TA630).[8]

As lenvatinib or sorafenib can only be used first-line in RAI refractory or ineligible patients, and selpercatinib is recommended only within the CDF, the only relevant comparator for cabozantinib is best supportive care (BSC).

Table 1: NICE decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with locally advanced or metastatic differentiated thyroid carcinoma, whose disease is refractory to, or who are unsuitable for radioactive iodine, and whose disease has progressed during or after prior systemic therapy.	Adults with locally advanced or metastatic differentiated thyroid carcinoma, whose disease is refractory to, or who are unsuitable for radioactive iodine, and whose disease has progressed during or after prior systemic therapy.	N/A
Intervention	Cabozantinib (CABOMETYX®)	Cabozantinib (CABOMETYX®)	N/A
Comparator(s)	Best Supportive Care (BSC)	Best Supportive Care (BSC)	As per the final scope, BSC is the comparator. There are no other treatments recommended post first-line systemic treatment for RAI refractory DTC patients by NICE, NHSE or ESMO. ESMO does state that 'cabozantinib and lenvatinib [are] two potential choices for second-line treatment of patients who progress on sorafenib'. However, as described earlier, the sequence of treatment should be determined on each patient's response and ESMO cannot create an optimal sequence for advanced/metastatic DTC due to limited current evidence.[3]
Outcomes	Draft Scope:	Co-primary endpoints	N/A

	• overall survival • progression-free survival • response rate • adverse effects of treatment • health-related quality of life	• Objective response rate (confirmed per RECIST v1.1) • Progression-free survival Additional endpoints • Overall survival • Adverse effects of treatment • Health-related quality of life	
Impact of the technology beyond direct health benefits, and on the delivery of the specialised service	No special considerations stated in the final scope.	No special considerations stated in the final scope.	N/A
Special considerations including issues related to equity or equality	No special considerations stated in the final scope.	No special considerations stated in the final scope.	Further to the company's decision problem and final scope, we believe that special considerations should be made regarding the female prevalence of DTC. Females are much more likely to be diagnosed with thyroid cancer making up 72% of thyroid cancer cases in the UK. In England, the AS incidence rate for thyroid cancer in females is 8.7 and for male it is 3.6 per 100,000, respectively, a clear difference in the incidence between females and males.[9]

Abbreviations: BSC – Best supportive care; DTC – Differentiated thyroid cancer; ESMO – European Society for Medical Oncology; N/A – Not applicable; NHSE – National Health Service England; NICE – National institute for Health and Care Excellence; RAI – Radioactive iodine.

B.1.2. Description of the technology being evaluated

A summary of the technology being appraised in this submission is given in Table 2. The summary of product characteristics (SmPC) and the European public assessment report (EPAR) are provided in Appendix C.

Table 2: Technology being evaluated

UK approved name and brand name	Cabozantinib (Cabometyx®)
Mechanism of action	Cabozantinib is an oral multi-targeted inhibitor of RTKs, inhibiting several RTKs known to influence tumour growth, angiogenesis and cancer cell invasion or metastasis, including VEGFR2, RET, MET and AXL.[10–13]
Marketing authorisation/CE mark status	An application for the marketing authorisation for cabozantinib in this indication was submitted to the EMA on 27th July 2021, with the European Centralised decision (considered as final approval) received on 29th April 2022. The EC decision was provided to the MHRA to facilitate the recognition route, using the EMA approval. Ipsen received GB approval for the Type II extension of the indication in DTC for Cabometyx dated 10th May 2022 from the MHRA.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The indication is as follows: "CABOMETYX is indicated as monotherapy for the treatment of adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed during or after prior systemic therapy." See Appendix C for the Summary of Product Characteristics[1] and EPAR.[14]
Method of administration and dosage	Oral administration: One 60 mg tablet to be taken once daily. Management of suspected adverse drug reactions may require temporary treatment interruption and/or drug reduction of cabozantinib therapy. When dose reduction is necessary in monotherapy, it is recommended to reduce to 40 mg daily and then to 20 mg daily. Dose reductions are recommended for events that, if persistent, could become serious or intolerable. If a patient misses a dose, the missed dose should not be taken if it less than 12 hours before the next dose.
Additional tests or investigations	It is recommended to perform liver function tests (ALT, AST and bilirubin) before cabozantinib treatment and to monitor closely during treatment. Platelet levels should be monitored during cabozantinib treatment, and the dose modified according to the severity of the thrombocytopenia. All patients should be monitored for hypertension and treated as needed with standard anti-hypertensive therapy. Urine protein should be monitored regularly during cabozantinib treatment. When using cabozantinib, periodic monitoring with on-treatment ECGs and electrolytes (serum calcium, potassium, and magnesium) should

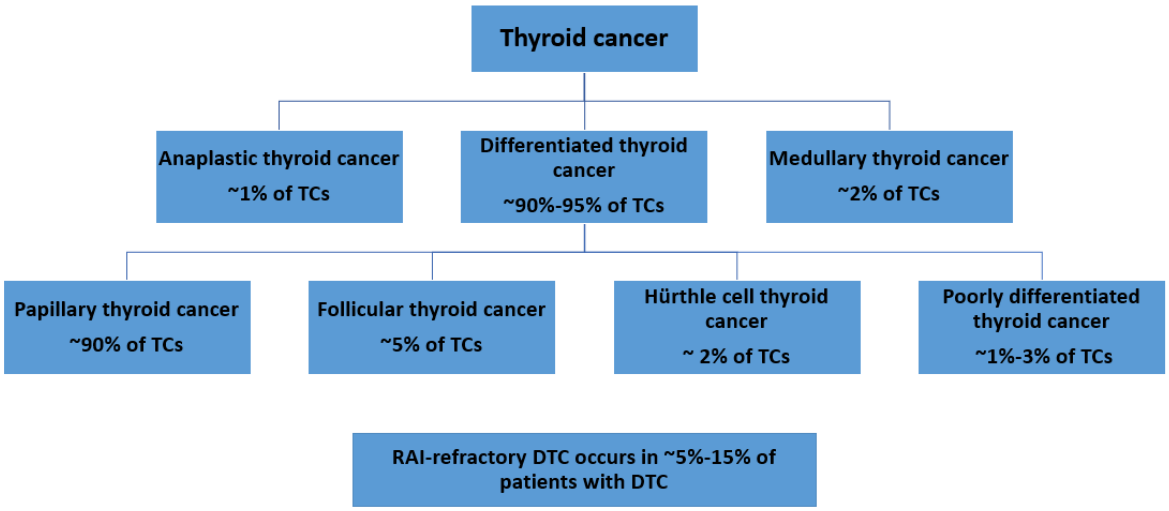
	be considered. Thyroid function should be monitored periodically throughout treatment with cabozantinib.[1]
List price and average cost of a course of treatment	£5,143.00 per 30 tablet pack.[15] The annual cost of cabozantinib at list price is £61,716.
Patient access scheme (if applicable)	A confidential simple patient access scheme is available. The pack price under this scheme is [redacted] (a [redacted]). The annual cost of treatment under this scheme is [redacted].

Abbreviations: ALT – Alanine transaminase; AST – Aspartate transaminase; DTC – Differentiated thyroid carcinoma; ECG – Electrogram; EMA – European Medicines Agency; EPAR – European Public Assessment Report; MET – Mesenchymal epithelial transition; MHRA – Medicines and Healthcare products Regulatory Agency; RAI – Radioactive iodine; RTK – Receptor tyrosine kinases.

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

Thyroid cancer (ICD-10-CM diagnosis code C73)[16] is a rare type of cancer that affects the thyroid gland. Thyroid cancers can be differentiated or undifferentiated, with DTC cells retaining the appearance of normal thyroid cells and usually growing slowly.[17] There are 4 main types of thyroid cancer: papillary, follicular, medullary and anaplastic. Two common types of DTC, papillary and follicular carcinomas, have similar management and prognosis. Figure 1 illustrates the distribution of thyroid cancer based on histology, with DTC being the most common form of thyroid cancer and accounting for ~90%-95% of all diagnosed cases.[18–20]

Figure 1: Classification of thyroid malignancies



Abbreviations: DTC – Differentiated thyroid cancer; RAI – Radioactive iodine; TC – Thyroid cancer.
Note: Proportions do not add up to 100% due to the ranges reported across multiple sources.
Sources: Rossi et al., 2021[21]; Miranda-Filho et al., 2021[22]; Lirov et al., 2017[18]; Gild et al., 2018[23]; Tumino et al., 2017[20]; Xu et al., 2020[24]

Thyroid cancer is uncommon and accounted for 1.2% of all new cases of cancer in the UK in 2020. There was a 5-year prevalence of 21,306 people with thyroid cancer in the UK in 2020.[25] DTCs are the most common types of thyroid cancers, with papillary carcinomas responsible for 90% of cases[17]. DTCs are typically curable, with 10-year survival typically around 85%.[26] Survival for thyroid cancer is strongly related to stage of disease. Survival is highest for adults diagnosed when the cancer is localised to the thyroid (Stage 1 to Stage 3), with 1-year age-standardised survival of around 99%.

Once the cancer has spread beyond the thyroid (Stage 4), 1-year age-standardised survival for adults diagnosed is 77%.[27]

Survival rates for RAI-refractory DTC is uncertain and is dependent on the availability of systemic therapies and prognosis of patients. For RAI-refractory DTC, the 5-year, 10-year and 15-year survival rates are 66%, 10% and 6% respectively.[18,28,29] Mortality rates become much worse for patients following progression from first line therapy (lenvatinib or sorafenib) if no salvage therapy is received. Studies have shown that median overall survival (OS) of patients who did not receive salvage therapy after progressing from a single agent TKI ranged between 10 months and 22 months.[30,31]

DTC usually has a good prognosis when treated with surgery, thyroid-stimulating hormone (TSH) suppression or RAI, used to destroy any remaining cancer cells.[2,32–35] External beam radiotherapy or palliative chemotherapy can also be used. The 2014 British Thyroid Association's 'Guidelines for the management of thyroid cancer' notes that the use of external beam radiotherapy and chemotherapy in palliative care has begun to be superseded by targeted therapy.[36] In clinical practice, BSC is offered until the disease starts to progress and symptoms occur, or there is rapid progression that is likely to become symptomatic.

For residual or recurrent disease, targeted therapy (TKIs) may be used. NICE TA535 recommends lenvatinib and sorafenib, which inhibit multiple receptor tyrosine kinases (RTKs) including VEGFRs, as options for treating DTC after RAI.[4] NICE TA742 recommends selpercatinib for use within the CDF, as an option for treating advanced RET fusion-positive thyroid cancer in adults who need systemic therapy after sorafenib or lenvatinib.[6] A number of guidelines have also been published over the last decade for the diagnosis and treatment of thyroid cancers, including the National Comprehensive Cancer Network (NCCN),[37] ESMO[3] and European Thyroid Association (ETA)[38].

ESMO states that lenvatinib and sorafenib should be considered the standard first-line systemic treatment for RAI-refractory DTC[3], with NCCN[39] preferring lenvatinib to sorafenib, however also stating both should be considered for patients with progressive and/or symptomatic disease. However, there is limited guidance on

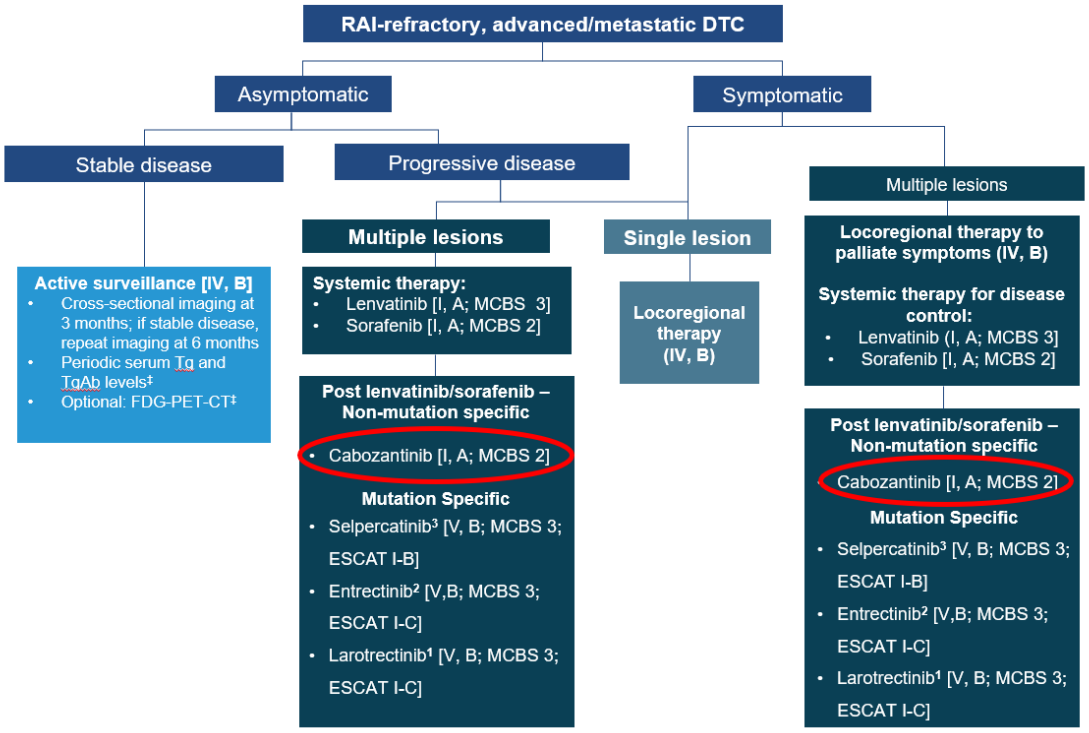
second-line and subsequent treatment for RAI-refractory DTC. In April 2022, ESMO released an update to their guidance on the use of systemic therapy in advanced thyroid cancer. The guideline states that cabozantinib is a potential choice for second-line treatment, but that the optimal sequence cannot be determined based on currently available evidence.[3]

The last decade has seen substantial research and development in novel targeted agents to treat patients with RAI-refractory DTC, however, there still remains a substantial unmet medical need for these patients. The standard of care of RAI-refractory DTC is systemic treatment with TKI, as discussed previously, with ESMO recommending lenvatinib and sorafenib.[40–42] Larotrectinib has been approved for NTRK fusion-positive solid tumours in adults and children, this includes thyroid cancer. Similarly, entrectinib is approved for NTRK fusion-positive solid tumours in persons over 12 years of age. Although, NTRK fusion-positive tumours are very rare in DTC.[43]

Larotrectinib (TA630) and entrectinib (TA644) are potential first or second-line treatments in this indication. NICE recommends that these treatments should be considered if patients “have no satisfactory treatment options”. In thyroid cancer there was an acceptance from both company and NICE that positioning for these treatments was uncertain, but it would likely be in the second or subsequent line setting in thyroid cancer.[7,8]

Cabozantinib works in a similar way to lenvatinib and sorafenib, by binding in a reversible manner to a region of the kinase domain and inhibiting catalytic activity, preventing further proliferation of the cancer.[13] Pre-clinical studies have demonstrated that cabozantinib is a potent inhibitor of mesenchymal epithelial transition (MET), growth arrest-specific protein 6 receptor (AXL), RET and VEGFR2, all of which are known to be important in the pathogenesis of thyroid cancer, specifically DTC.[11–13,44,45] The simultaneous targeting of these pathways by cabozantinib may provide enhanced anti-tumour effects compared to agents that target only one of these pathways.[13] Figure 2 outlines the potential positioning of cabozantinib within the treatment pathway based on ESMO guidelines and specified to the UK from NICE recommendations.

Figure 2: DTC Treatment overview for RAI-refractory DTC, including cabozantinib (adapted from ESMO and NICE recommendations)



Abbreviations: DTC – Differentiated thyroid cancer; EMA – European Medicines Agency; ESCAT – ESMO Scale for Clinical Actionability of molecular Targets; ESMO-MCBS – ESMO-Magnitude of Clinical Benefit Scale; FDG-PET – [18F]2-fluoro-2-deoxy-D-glucose-positron emission tomography; FDG-PET-CT – [18F]2-fluoro-2-deoxy-D-glucose-positron emission tomography-computed tomography; MCBS – ESMO-Magnitude of Clinical Benefit Scale; MKI – Multikinase inhibitor; NGS – Next-generation sequencing; NTRK – Neurotrophic tyrosine receptor kinase; RAI – Radioactive iodine; RECIST – Response Evaluation Criteria in Solid Tumours; TC – Thyroid cancer; Tg – Thyroglobulin; TgAb – Serum thyroglobulin antibody.
Source: Filetti S, et al. 2022 Ann Oncl; 33: 674-684,[3] NICE TA (535, 630, 644, 742)[4,6–8]

B.1.4. Equality considerations

Thyroid cancer European age standardised (AS) incidence rates for females and males combined, increased by 175% in the UK between 1993-1995 and 2016-2018. The increase was of a similar size in females and males. Females are much more likely to be diagnosed with thyroid cancer making up 72% of thyroid cancer cases in the UK. The AS incidence for thyroid cancer in females is 8.7 and for male it is 3.6 per 100,000, respectively. Therefore, cabozantinib in DTC will reduce the health inequalities for female thyroid cancer patients.[9]

B.2. Clinical effectiveness

B.2.1. Identification and selection of relevant studies

A systematic literature review (SLR) was undertaken on the 14th October 2021 to identify published clinical studies relevant to the decision problem (see Section B.1.1).

Please see Appendix G for the methods used to identify all relevant studies.

B.2.2. List of relevant clinical effectiveness evidence

Clinical evidence to support the use of cabozantinib in adult patients with locally advanced or metastatic DTC that have progressed following prior VEGFR-targeted therapy and who are ineligible, or RAI-refractory, comprises a single randomised control trial (RCT) – the COSMIC 311 trial (XL184-311; NCT03690388). A brief overview of this trial is outlined in Table 3.

A SLR did not identify any additional studies relevant to cabozantinib in RAI-refractory advanced DTC. No network meta-analysis (NMA) was needed as currently the only relevant comparator in England and Wales for this population is BSC which is informed by matched placebo control arm of the COSMIC-311 trial.

Table 3: Clinical effectiveness evidence for COSMIC-311

Study	COSMIC - 311
Study design	Randomised, double-blind, placebo-controlled, phase III
Population	Patients with previously treated advanced RAI-Refractory DTC
Intervention(s)	Oral cabozantinib 60 mg once daily plus best supportive care (BSC)
Comparator(s)	Oral matched placebo once daily plus BSC
Indicate if study supports application for marketing authorisation	Yes
Indicate if study used in the economic model	Yes
Reported outcomes specified in the decision problem	• Overall survival (OS) • Progression-free survival (PFS) • Time to treatment discontinuation (TTD) • Objective response rate (ORR) • Adverse events (AEs) • Health-related quality of life (EQ-5D-5L)
All other reported outcomes	Pharmacokinetics

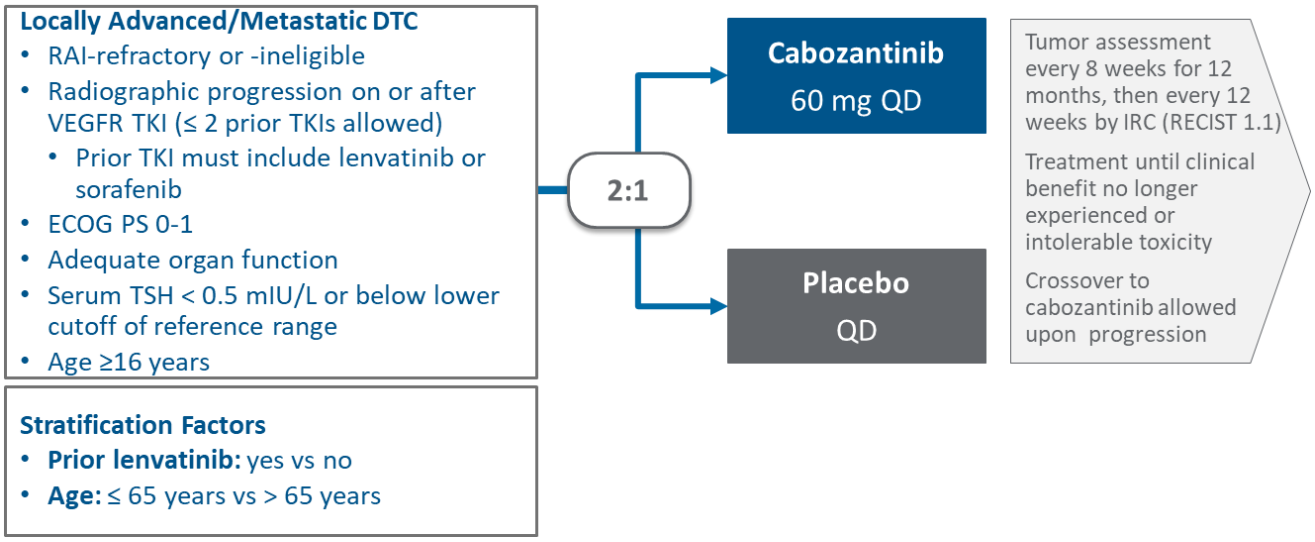
Abbreviations: AE – Adverse events; BSC – Best supportive care; DTC – Differentiated thyroid carcinoma; EQ5D-5L – Health-related quality of life; ORR – Objective response rate; OS – Overall survival; PFS – Progression-free survival; TTD – Time to treatment discontinuation.

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

B.2.3.1 COSMIC-311 Trial

The COSMIC-311 global phase III clinical trial tested the efficacy and safety of cabozantinib in adult patients with RAI-refractory advanced DTC, who have progressed during or after previous systemic therapy. The COSMIC-311 trial schematic design is outlined in Figure 3. Table 4 provides a summary of the trial methodology.

Figure 3: Trial design of COSMIC-311



Abbreviations: ECOG – Eastern Cooperative Oncology Group; IRC – Independent Radiology Committee; QD – Once a day; RAI – Radioactive iodine; RECIST – Response Evaluation Criteria in Solid Tumours; TKI – Tyrosine kinase inhibitor; TSH – Thyroid-Stimulating Hormone; VEGFR – Vascular endothelial growth factor receptor.
Source: XL184-311 CSR (30th April 2020)[46]

Table 4: Summary of trial methodology

Trial	COSMIC-311
Trial design	Phase 3, randomised, multicentre, double-blind, 2:1 controlled study of cabozantinib versus placebo in patients with RAI-refractory DTC who have received prior lenvatinib or sorafenib treatment.

Eligibility criteria for participants	**Key Inclusion criteria:**
	• Histologically or cytologically confirmed diagnosis of DTC, including the following subtypes:
	  ○ PTC including histological variants of PTC
	  ○ FTC including histological variants of FTC
	• Measurable disease according to RECIST 1.1 on CT/MRI performed within 28 days prior to randomisation
	• Must have been previously treated with or deemed ineligible for treatment with Iodine-131 for DTC
	• Patients must have received at least one prior VEGFR-targeting TKI therapy of either lenvatinib or sorafenib and must have had radiographic progression during treatment or within 6 months after the most recent dose of the VEGFR inhibitor (up to two prior therapies were allowed including, but not limited to, lenvatinib and sorafenib)
	• Must have experienced documented radiographic progression per RECIST 1.1 per the Investigator during or following treatment with a VEGFR-targeting TKI prior to starting the next anticancer therapy (which may have been treatment in COSMIC-311)
	• Age – 16 years and older (Adult, Older Adult)
	• ECOG PS of 0 or 1

Exclusion Criteria for participants	**Key Exclusion criteria:**
	• Prior treatment with any of the following:
	  ○ Cabozantinib
	  ○ Selective small-molecule BRAF kinase inhibitor (e.g., vemurafenib, dabrafenib)
	  ○ More than 2 VEGFR-targeting TKI agents (e.g., lenvatinib, sorafenib, sunitinib, pazopanib, axitinib, vandetanib)
	  ○ More than 1 immune checkpoint inhibitor therapy (e.g., PD-1 or PD-L1 targeting agent)
	  ○ More than 1 systemic chemotherapy regimen (given as single agent or in combination with another chemotherapy agent)
	• Receipt of any type of small molecule kinase inhibitor within 2 weeks or 5 half-lives of the agent, whichever was longer, before randomisation

	• Receipt of any type of anticancer antibody or systemic chemotherapy within 4 weeks before randomisation • Receipt of radiation therapy for bone metastasis within 2 weeks or any other radiation therapy within 4 weeks before randomisation • Subjects with clinically relevant ongoing from prior radiation therapy that had not completely resolved were not eligible *All inclusion and exclusion criteria are listed in appendix C.*
Settings and locations where the data were collected	A total of 258 subjects were randomised in 161 unique sites by 174 principal investigators in 25 countries in Asia, North America, Europe, and the rest of the world. These included: • Europe: Austria, Belgium, Croatia, Czech Republic, France, Germany, Hungary, Italy, Netherlands, Poland, Romania, Spain, United Kingdom • North America: United States of America and Canada • Asia: Hong Kong, Republic of Korea, Taiwan, Thailand • Rest of the world: Argentina, Australia, Brazil, Israel, Mexico, Russia,
Trial drugs	• Experimental Arm: Cabozantinib 60 mg tablet once daily ○ Two dose reductions in decrements of 20 mg was permitted to manage or prevent AE or toxicity • Comparator Arm: Matched placebo
Permitted and disallowed concomitant medication	**Allowed concomitant medication** • Prophylactic antiemetics and antidiarrheal medications in line with standard clinical practice • Granulocyte colony-stimulating factors per ASCO or ESMO guidelines • Bisphosphonates or denosumab for the control of bone loss or hypercalcemia if the benefit per the Investigator's discretion • Transfusions and hormone replacement (including TSH-suppressive thyroid hormone therapy) • Prophylactic individualised anticoagulation therapy with low dose low molecular weight (LMWH) heparins for supportive treatment per the Investigator's discretion. LMWH use at first dose should only be used if the subject had no evidence of brain metastasis, had been on stable dose of LMWH for a least six weeks prior, and had no complications from a thromboembolic event or the anticoagulation regimen. Therapeutic doses of oral anticoagulants (e.g., warfarin or other coumarin-related agents) were not allowed after randomisation until study treatment was permanently discontinued **Prohibited Therapies**

	• Any investigational agent or investigational medical device • Any systemic NPACT (e.g., chemotherapy, immunotherapy, radionuclides, drugs, or herbal products used specifically for the treatment of DTC). • Therapeutic doses of oral anticoagulants. • Local anticancer treatment including palliative radiation, ablation, embolisation or surgery impacting on tumour lesions were only to be performed until radiographic progression was confirmed per RECIST 1.1. • Erythropoietic-stimulating agents prohibited due to the increased risk of tumour recurrence. • Concomitant medications that prolong the QTc interval were to be avoided until subjects discontinue treatment. • Chronic coadministration of strong CYP 3A4 inducers due to potential to decrease exposure to cabozantinib. • Coadministration of strong CYP3A4 inhibitors and other drugs that inhibit CYP3A4 was to be avoided because these drugs had the potential to increase exposure (AUC) to cabozantinib

Abbreviations: AE – Adverse event; AUC – Area under the curve; ASCO – American Society of Clinical Oncology; CT – Computed tomography; CYP – Cytochrome P450; DTC – Differentiated thyroid cancer; ECOG – Eastern Cooperative Oncology Group; ESMO – European Society of Medical Oncology; FTC – Follicular thyroid carcinoma; LMWH - Low molecular weight heparin; MRI – Magnetic resonance imaging; NPACT – Nonprotocol anticancer therapy; PD-1 – Programmed cell death-1; PD-L1 – Programmed cell death ligand 1; PS – Performance status; PTC – Papillary thyroid carcinoma; QTc – Corrected QT interval; RAI – Radioactive iodine; RECIST – Response Evaluation Criteria in Solid Tumours; TKI – Tyrosine kinase inhibitor; VEGFR – Vascular endothelial growth factor receptor.
Source: XL184-311 CSR (30th April 2020)[46] and Brose et al, 2021[47]

B.2.3.1.1. COSMIC-311 endpoints

A total of 258 subjects were randomised 2:1 to receive either cabozantinib or placebo, with 177 receiving cabozantinib and 88 receiving placebo. The co-primary endpoints were objective response rate (ORR) and progression free survival (PFS).

The primary analysis of ORR was limited to the first 100 randomised subjects and was defined as the proportion of patients with a best overall response (BOR) of confirmed complete response (CR) or confirmed partial response (PR) per RECIST 1.1 by blinded independent radiology committee (BIRC). PFS was defined as the time from randomisation until progressed disease (PD) or death. An event in the PFS analysis was determined by radiographic progression as determined by BIRC per RECIST 1.1 and death. Secondary endpoints included: OS, duration of objective response (DOR), time to objective response and health-related quality of life (HRQoL).

Data from two data cuts are available: 19th August 2020 and 8th February 2021 as clinical cut-off 1 (CCO1) and clinical cut-off 2 (CCO2), respectively. Table 5 provides an outline of all primary and secondary endpoints and the data cut available for analysis.

Table 5: Relevant endpoints definitions and measures in the COSMIC-311 trial

Endpoint	Definition	Timing and nature of assessment	Clinical Cut-off available
Primary endpoints			
ORR	The proportion of patients with a BOR of CR or PR. CR or PR must be confirmed on a subsequent visit ≥28 days after the response was first observed	The first 100 randomised subjects were followed up for the primary analysis of ORR. CT/MRI assessment of the chest, abdomen and pelvis were performed at screening, 8 weeks after randomisation and every 8 weeks thereafter. CT/MRI of the brain was performed at screening and as clinically indicated (suspicion of brain metastases). Whole-body bone scans were acquired for all subjects at screening using a technetium-99 (99Tc) bone seeking radiopharmaceutical; follow-up scans were performed every 24 weeks (± 14 days) thereafter only for subjects who had documented bone metastases. Assessments continued until 8 weeks after investigator-defined radiographical disease progression or the date of the decision to permanently discontinue study drug, whichever came first, irrespective of whether study drug was given or the dose was reduced, interrupted, or discontinued. After the post-treatment follow-up visit 30 days after the decision to discontinue study drug, patients were contacted every 8 weeks to assess their survival status	CCO1 CCO2
PFS	The date of randomisation to radiographical progression as determined by BIRC per RECIST 1.1 or death, whichever occurred first	Primary analysis of PFS included radiographic progression events as determined by BIRC per RECIST 1.1 and deaths. Clinical deterioration or radiographic progression determined by the Investigator were not to be considered as events for the primary analysis.	CCO1 CCO2
Secondary endpoints			
OS	The date of randomisation until death due to any cause.	After the post-treatment follow-up visit 30 days after the decision to discontinue study drug, patients were contacted every 8 weeks to assess their survival status.	CCO1 CCO2
Exploratory endpoints			
HRQoL	Health status was measured using EQ-5D-5L		CCO1

	The EQ-5D-5L questionnaire was self-administered by the patient at baseline, every 4 weeks for 25 weeks and every 8 weeks thereafter, regardless of whether study drug was given, or the dose was reduced, interrupted, or discontinued, until 8 weeks after either disease progression according to RECIST 1.1 or the decision to permanently discontinue study drug. The EQ-5D-5L questionnaire was not given to patients who spoke a language for which there was not an approved translation of the questionnaire.	CCO2
Safety and tolerability	Safety assessments included the evaluation of AEs, SAEs, deaths, clinical laboratory tests (haematology, serum chemistry and urinalysis), physical examination, vital signs, ECOG PS, 12-lead ECG and the TTD in months (date of decision to discontinue study drug – date of first dose +1)/30.4375. Safety was monitored throughout the trial. Safety was assessed at least every 2 weeks for the first 9 weeks, then every 4 weeks thereafter, irrespective of any dose interruptions, with the final assessment 30 days after the decision to discontinue study drug (unless there was an ongoing Grade 3 or 4 AE or SAE) The severity of AEs, whether they were SAEs and their potential relationship to study drug were assessed by the investigator. Severity was defined by Common Terminology Criteria for Adverse Events (CTCAE) version 4. The Safety Committee and an Independent Data Monitoring Committee (IDMC) monitored safety on a regular basis.	CCO1 CCO2
Pre-planned subgroups	Exploratory analysis of PFS, ORR, and OS were conducted to evaluate the effect of subgroups based on baseline characteristics. These included: • Age (≤ 65 years, > 65 years) • Sex (Male, Female) • Race (Asian, Black, White, other/not reported) • Prior sorafenib or lenvatinib therapy, or both • Prior VEGFR-TKI anticancer therapy agents • ECOG PS at baseline (0, 1) • Histology (Papillary, Follicular) • Bone, important visceral, liver, lung, metastases per Investigator (Yes, No)	CCO1 CCO2

Abbreviations: AE – Adverse event; AUC – Area under the curve; ASCO – American Society of Clinical Oncology; CT – Computed tomography; CYP – Cytochrome P450; DTC – Differentiated thyroid cancer; ECOG – Eastern Cooperative Oncology Group; ESMO – European Society of Medical Oncology; FTC – Follicular thyroid carcinoma; LMWH - Low molecular weight heparin; MRI – Magnetic resonance imaging; NPACT – Nonprotocol anticancer therapy; PD-1 – Programmed cell death-1; PD-L1 – Programmed cell death ligand 1; PS – Performance status; PTC – Papillary thyroid carcinoma; QTc – Corrected QT interval; RAI – Radioactive iodine; RECIST – Response Evaluation Criteria in Solid Tumours; TKI – Tyrosine kinase inhibitor; VEGFR – Vascular endothelial growth factor receptor.
Source: XL184-311 CSR (30th April 2020)[46]

B.2.3.2 Crossover phase in COSMIC-311

A feature of the trial design for COSMIC-311 was the permitting of crossover of subjects in the placebo arm to the cabozantinib arm upon radiographic PD per RECIST 1.1 and confirmed by the BIRC. The subjects that crossed treatment arms were subsequentially unblinded and were to continue on treatment if the Investigator believed the subject was still receiving clinical benefit. Subjects who crossed over were to continue safety assessments and tumour assessments as per the study protocol, however these scans were not submitted to BIRC. Also, pharmacokinetics (PK) and HRQoL assessments were discontinued for these subjects. Section B.2.4.2 outlines the methods of adjusting endpoint results for crossover in the trial.

A comprehensive outline of the eligibility of treatment crossover can be found in the COSMIC-311 study protocol.[48]

B.2.3.3 Patient baseline demographics and disease characteristics

The baseline characteristics of the Full intention to treat (ITT) population from the latest data cut-off date on the 8th of February 2021 (CCO2) is shown in Table 6. The total number of patients recruited in to COSMIC-311 was 258 with 170 in the cabozantinib arm and 88 in the placebo arm.

The demographic and baseline characteristics are representative of DTC epidemiology, with the median age at 65 years for the cabozantinib arm and 66 years for placebo. Approximately 50% of patients were 65 years of age or younger and there were slightly more female (53%) patients than male (47%). Patients from Europe made up 47% of the patient population in the study. The proportion of patient's refractory to prior RAI therapy was balanced between treatment arms and were similar in the populations. Prior non-radiation anticancer therapies were also well balanced between treatment arms and similar across the populations. The majority of patients had received only one prior VEGFR-TKI. In the Full ITT, approximately 63% of patients had received prior treatment with lenvatinib and 60% had received prior treatment with sorafenib, with 24% having received prior treatment with both.

Overall, the baseline demographics and disease characteristics were balanced between treatment arms and were similar across the ORR ITT (OITT), ITT and Full ITT population. See Section B.2.4.1 below for the details regarding these population definitions.

Based on clinical advice it is expected that the baseline characteristics of COSMIC-311 are representative of the patient population in England.[49,50] A subset of patients in this study received both sorafenib and lenvatinib before receiving the study treatment (24%). It is against NICE guidance (TA535) to receive both lenvatinib and sorafenib, with only one used as first line therapy, and neither are approved for second-line in England and Wales.[4] Seventy-four per cent of patients in the ITT population in both the cabozantinib and placebo arms only received one VEGFR inhibitor (sorafenib or lenvatinib or other), reflecting the potential population for cabozantinib in England and Wales based on NICE guidance.[48]

Table 6: Characteristics of participants in the studies across treatment groups (CCO2)

COSMIC-311 **Baseline characteristic**	**Cabozantinib**	**Placebo**
Full ITT population	n=170	N=88
Age, median years (range) **≥ 65 years (%)**	65 (31-85) [redacted]	66 (37-83) [redacted]
Sex n (%) Male Female	 83 (49) 87 (51)	 39 (44) 49 (56)
Geographical Region n (%) Europe Asia North America (USA and Canada) Rest of the world	 82 (48) 24 (14) 15 (8.8) 49 (29)	 39 (44) 19 (22) 12 (14) 18 (20)
Race, n (%) White Asian Black or African American Other / Not reported	[redacted]121 (71) 29 (17) 2 (1.2) 18 (10.6)	[redacted]59 (67) 20 (23) 2 (2.3) 7 (7.9)
ECOG PS, n (%) 0 (normal activity, asymptomatic) 1 (fully ambulatory, symptomatic)	 74 (44) 96 (56)	 43 (49) 45 (51)
Smoking history, n (%) Current Former	[redacted]	[redacted]

Never		
Weight, median (range) (kg)	[redacted]	[redacted]
BMI, median (range) (kg/m2	[redacted]	[redacted]
Previous sorafenib or lenvatinib n (%) Sorafenib but no lenvatinib Lenvatinib but no sorafenib Sorafenib and lenvatinib Other TKI therapy	 61 (36) 68 (40) 40 (23) 1	 33 (38) 34 (39) 21 (24) 0
Number of previous vascular endothelial growth factor receptor tyrosine kinase inhibitors n (%) 0 1 2	 1 126 (74) 43 (25)	 0 65 (74) 23 (26)
Histological subtype n (%) 1 Papillary Follicular	 96 (56) 78 (46)	 54 (61) 35 (43)
Metastatic lesions n (%) Bone Liver Lung Other	159 (94) 51 (30) 25 (15) 121 (71) 127 (75)	82 (93) 21 (24) 9 (10) 61 (69) 70 (80)
Number of prior PD-1/PD-L1 agents per subject for DTC, n (%) 0 1 ≥2 Median (range)	[redacted]	[redacted]
Median (range) time from progression on most recent prior non-radiation systemic anticancer regimen for DTC to randomisation, months	[redacted]	[redacted]

1: Patients could be counted as having both Papillary and Follicular histological subtypes.
Abbreviations: ITT — Intention to treat population, USA — United Sates of America, ECOG — Eastern Cooperative Oncology Group, PS — Performance status, PD-1 — Programmed cell death protein 1, PD-L1 — Programmed death ligand
Source: XL184-311 CSR Addendum 1 (21st May 2021)[51]

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

B.2.4.1 Analysis Population

Approximately 300 subjects with RAI-refractory DTC were planned to be randomised to receive study treatment. The OITT population consisted of the first 100 subjects who were randomised. The COSMIC-311 study employed a "trial within a trial design"[52], where the first 100 patients who were randomly assigned were considered

a subpopulation for analysis of the ORR. This was designed to obtain an earlier evaluation of ORR, with the first interim analysis reporting results of PFS.

The primary data cut-off (CCO1) was 19th August 2020 and the supportive efficacy analysis using the follow-up data cut-off was 8th February 2021 (CCO2).

Through to the 19th August 2020 cut-off date (CCO1) 187 subjects (125 cabozantinib, 62 placebo) were randomised to receive study treatment (ITT population).

After the 19th August 2020 cut-off date, subjects continued to enroll in the study and receive blinded study treatment. Given that the study demonstrated significant improvement in PFS at the interim analysis, enrollment was stopped and the last subject was randomised on 2nd February 2021. A total of 258 subjects (170 cabozantinib, 88 placebo) were enrolled in the study. The second data cut-off date was 08 February 2021 (CCO2). The supportive analyses were performed on the Full ITT population, defined as all 258 subjects randomised to the study as of CCO2.

See Table 7 for an overview of the patient numbers of the CCO1, CCO2 and safety populations. The safety population comprised all patients who were randomised to receive and received at least one dose of study drug (cabozantinib or matched placebo).

Table 7: Analysis population for the COSMIC-311 trial

Analysis Populations	**Number of Patients**		
	Cabozantinib	**Placebo**	**Total**
CCO1 (19th August 2020)			
ITT population **Safety**	125	62	187
CCO2 (8th February 2021)			
Full ITT population **Safety**	170	88	258

Abbreviations: CCO – Clinical cut-off; ITT — Intention to treat.

NOTE: There are three Clinical Study Reports COSMIC-311 related to the above populations as listed below.

Analysis Population	CSR filename (pdf) and date	Comment
CCO1 (Efficacy and Safety)	xl184-311-csr (CCO1) CSR date: 30th April 2020	Efficacy, HRQoL and safety data
CCO2 (Full ITT)	xl184311csr-body-addendum-1 (CC02) CSR date 21st May 2021	Efficacy data only
CCO2 (Safety)	xl184311 csr body addendum 2 (CC02) CSR Date: 19th August 2021	Safety data only

B.2.4.2 Statistical Analysis

The primary efficacy analyses for this study compared the results in subjects randomised to receive cabozantinib to those in the placebo arm for the multiple primary endpoints ORR and PFS. Treatment with cabozantinib would be inferred to be superior to treatment with placebo if the null hypothesis of no difference between arms was rejected for either ORR or PFS.

Analysis of the additional endpoint OS was descriptive and non-inferential as OS was not a controlled endpoint for the study. The primary purpose of the OS analyses was to evaluate the potential for detriment to survival with cabozantinib treatment.

Table 8 summarises the methods used for the statistical analysis.

Table 8: Summary of the statistical analyses undertaken in the COSMIC-311 trial

Hypothesis objective	Statistical analysis	Sample size, power calculation
The null hypothesis was that there was no difference in the duration of PFS between the treatment groups (cabozantinib plus BSC versus placebo plus BSC) The alternative hypothesis was that there was a difference in the duration of PFS between the treatment groups (cabozantinib plus BSC versus BSC)	**Primary efficacy analyses** *Primary efficacy endpoint; PFS and ORR* Analyses: A single interim analysis of PFS was planned at the time of the primary analysis of ORR (19 August 2020). The primary analysis of PFS was event driven and was planned to be conducted when 193 events had been observed (radiographic progression according to RECIST 1.1 or death). Final cut-off was the 8th of February 2021. Hypothesis testing was performed using the stratified log-rank test with a two-sided α=0.04 or α=0.05. The stratification factors were the same as those used to stratify randomisation (stratified by receipt of prior lenvatinib and age ≤ 65 years). The median duration of PFS and the associated 96% or 95% CI for each treatment arm were estimated using the Kaplan-Meier method. The stratified hazard ratio (HR) and its 96% or 95% CI were estimated using a Cox proportional-hazard model with treatment group as the independent variable and stratified by the same randomisation stratification factors used for the log-rank test. The control of type 1 error arising from the multiplicity issue from the primary analyses of PFS and ORR were addressed by applying a modified Bonferroni procedure. ORR was tested at the 2-sided 1% interval and PFS was tested at the 2 sided 4% significance level. Inflation type 1 error due to the multiple analyses of PFS was to be controlled using a Lan-DeMets O'Brien alpha-spending function. At the time of the first interim analysis of PFS 74 (20% cabozantinib vs. 66% placebo) events had been observed. The HR, adjusted for stratification factors (per IxRS), was 0.22 (96% CI: 0.13, 0.36; stratified	Two samples were used to determine the main endpoints of ORR and PFS. The ORR ITT population had a planned sample size of 100 subjects and an actual of 100 subjects. The overall ITT has a planned sample of 300 subjects with an actual of 187. It was estimated that 193 events were to be observed in 300 patients in the ITT population would be needed to determine 90% power for a two-sided log-rank test

	log-rank p-value < 0.0001). The primary PFS was tested at the observed 38.3% information fraction using a critical p-value of 0.00036 and the null hypothesis was not rejected. If the p-value was less than the critical value for rejecting the null hypothesis and the HR was <1, the null hypothesis was rejected, and it was inferred that PFS was superior in the cabozantinib group compared with the placebo group. Results of the interim analyses were evaluated by the IDMC to allow the trial to be stopped early if the null hypothesis for PFS was rejected in favour of cabozantinib. ORR: The primary efficacy endpoint of ORR was defined as the proportion of subjects with a best overall response (BOR) of confirmed complete response (CR) or confirmed partial response (PR) per RECIST 1.1. by BIRC. The confirmation must have occurred at least 28 days after the response of CR or PR was observed. Hypothesis testing was performed using the Fisher's exact test at the 2-sided α=0.01 level of significance. Analysis using the Cochran-Mantel-Haenszel (CMH) method to adjust for stratification factors per IRT was also conducted. Point estimates of ORR, the difference in ORR between the two treatment arms, and associated CIs were provided. The odds ratio and its CIs were also provided. The two-sided 95% and 99% CIs were calculated using exact methods except for the difference in ORR between the two treatment arms and for the odds ratio which used asymptotic confidence limits. If the p-value for the two-sided Fisher's exact test was less than 0.01 and the point estimate for ORR in the cabozantinib arm was higher than that in the placebo arm, the null hypothesis of no difference in ORR was rejected, inferring that ORR was superior in the cabozantinib arm compared with the placebo arm.	at 5% significance to detect a 64% increase in PFS with cabozantinib compared with placebo (HR 0.61).

	Secondary efficacy endpoint OS: The trial design meant it was not possible to show powered OS results. The primary purpose of the OS analysis was to evaluate the potential for detriment to survival with cabozantinib treatment. The median duration of OS and the associated 95% CI for each treatment arm were estimated using the Kaplan-Meier method. The unstratified and stratified HR and their 95% CI were estimated using a Cox proportional-hazard model with treatment group as the independent variable. Log-rank p-values were calculated and presented for descriptive purposes; formal inferences were not drawn.	

Abbreviations: BOR – Best overall response; BSC – Best supportive care; CR – Complete response; CI – Confidence interval; IxRS – Interactive voice/web response system ORR – Objective response rate; OS – Overall survival; PFS – Progression-free survival; PR – Partial response
Source: XL184-311 CSR (30th April 2020)[46]

B.2.4.2.1. Sensitivity Analyses

Multiple sensitivity analyses of PFS were performed on the ITT population in CCO2. Sensitivity analyses in CCO2 were limited to radiographic progression based on BIRC only. Two sensitivity analyses were performed using different assumptions and interpretations of condition (tumour) assessment. Two other sensitivities were conducted using different criteria for determining whether a situation is considered an event or censored. All sensitivity analyses performed in CCO2 are as follows:

- PFS-EP-2: Evaluated the influence of potentially inconsistent tumour assessment intervals between arms. For subjects who experienced radiographic progression, it assigned the date of the scheduled visit as the event date, rather than the date of recorded progression.

- PFS-EP-4: Evaluated the influence of missing tumour assessments. It classified subjects who experienced ≥ 2 consecutive missing scheduled adequate tumour assessments (ATA) immediately prior to documented radiographic progression as having an event, rather than being censored, at the date of the last ATA prior to the missing visits.

- PFS-EA2-1: Receipt of systemic NPACT was changed to “composite,” resulting in an endpoint that comprised radiographic progression, death, or initiation of systemic NPACT

- PFS-EA2-2: Combined the sensitivity analysis of PFS-EP-4 and PFS-EAS2-1.

Table 9 outlines the censoring and event rules used in the primary analyses of PFS and the sensitivity analyses. For OS, patients who were alive at the time of data cut off or who were permanently lost to follow up, duration of OS was censored at the date the patient was last known to be alive.

Table 9: Event and censoring rules for the primary and sensitivity analysis of PFS

Analysis	Primary Analysis (PFS)		PFS-EP-2		PFS-EP-4		PFS-EA2-1		PFS-EA2-2	
Situation	**Outcome**	**Date**	**Outcome**	**Date**	**Outcome**	**Date**	**Outcome**	**Date**	**Outcome**	**Date**
Radiographic PD per RECIST 1.1 per BIRC	Event	Date of recorded PD	Event	Date of scheduled visit (or next scheduled visit if between visits)	Event	Date of recorded PD	Event	Date of recorded PD	Event	Date of recorded PD
Radiographic PD per RECIST 1.1 per Investigator	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Death	Event	Date of death	Event	Date of death	Event	Date of death	Event	Date of death	Event	Date of death
Clinical deterioration	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Systemic NPACT (medications)	Censored	Date of last ATA* before first initiation of therapy	Censored	Date of last ATA* before first initiation of therapy	Censored	Date of last ATA* before first initiation of therapy	Event	Date of first initiation of therapy	Event	Date of first initiation of therapy
Local NPACT (medications	Censored	Date of last ATA* before first initiation of therapy	Censored	Date of last ATA* before first initiation of therapy	Censored	Date of last ATA* before first initiation of therapy	Censored	Date of last ATA* before first initiation of therapy	Censored	Date of last ATA* before first initiation of therapy
Surgical resection of target tumour lesion (s)	Censored	Date of last ATA* before target	Censored	Date of last ATA* before target lesion resection	Censored	Date of last ATA* before target	Censored	Date of last ATA* before first	Censored	Date of last ATA* before first

		lesion resection				lesion resection		initiation of therapy		initiation of therapy
Local radiation: to soft tissue for disease under study	Censored	Date of last ATA* local radiation of soft tissue for disease under study	Censored	Date of last ATA* local radiation of soft tissue for disease under study	Censored	Date of last ATA* local radiation of soft tissue for disease under study	Censored	Date of last ATA* local radiation of soft tissue for disease under study	Censored	Date of last ATA* local radiation of soft tissue for disease under study
No baseline ATA	Censored	Date of rand.	Censored	Date of rand.	Censored	Date of rand.	Censored	Date of rand.	Censored	Date of rand.
≥ 2 consecutive missing scheduled ATA immediately prior to analysis even	Censored	Date of last ATA* before missing visits	Censored	Date of last ATA* before missing visits	Event	Date of last ATA* before missing visits	Censored	date of last ATA* before missing visits	Event	Date of last ATA* before missing visits
None of the above	Censored	Date of last ATA	Censored	Date of last ATA	Censored	Date of last ATA	Censored	Date of last ATA	Censored	Date of last ATA

Blue boxes indicate differing outcomes or dates of recording outcome. * Or date of randomisation if no post-randomisation ATA.
Abbreviations**:** ATA – Adequate tumour assessment; BIRC – Blinded independent review committee; ITT – Intent-to-treat; NA – Not applicable; NPACT – Non-protocol anti-cancer therapy (medications including radiopharmaceuticals but excluding local radiation); PD – Progressive disease; rand – Randomisation; RECIST – Response Evaluation Criteria in Solid Tumours; rPFS — radiographic progression-free survival.
Source: XL184-311 CSR (30th April 2020)[46]

B.2.4.2.2. COSMIC-311 Crossover Adjustment

If eligible and upon investigator request, patients receiving placebo treatment were allowed to crossover to cabozantinib treatment after disease progression. Patients that switched from placebo to cabozantinib treatment entered the 'crossover phase' upon treatment switch. Data from this period was collected independently.

Due to a significant level of crossover in the COSMIC-311, it is necessary to mitigate bias in the OS results by adjusting for crossover. Traditional ITT analysis is not appropriate for the analysis of the COSMIC-311 OS data as it does not account for the possible OS benefit received by placebo patients who switched to cabozantinib and can therefore underestimate the relative efficacy of cabozantinib compared to a true placebo arm that does not include patients receiving subsequent cabozantinib treatment. There are multiple statistical methods that can be used to adjust for treatment switching when analysing clinical trial data. However, they all come with important assumptions and limitations that need to be acknowledged. For instance, the relatively "simple" statistical adjustment methods such as censoring switchers at the point of switch or excluding them entirely from the analysis are highly prone to selection bias because switching is likely to be associated with differing prognosis.

The statistical analysis plan (SAP) outlined the potential for exploratory analysis to adjust for crossover in COSMIC-311 using the method of inverse probability of censoring weights (IPCW), the rank preserving structural failure time (RPSFT) or the "two-stage method", if any were deemed feasible. NICE technical support document (TSD) 16 gives a detailed description on the use of these methods.[53]

- IPCW: The IPCW method represents an observational-based approach for adjusting treatment effect estimates in the presence of any type of informative censoring. The IPCW method records observations in time intervals and artificially censors switchers at the point of treatment switch. A limitation of the IPCW method is that it relies on the "no unmeasured confounders" assumption, i.e., data must be available on all baseline and time-dependent prognostic factors for mortality that independently predict switching.

- RPSFT: RPSFT uses a counterfactual framework to estimate the survival time gained or lost by receiving active treatment, where counterfactual survival times refer to those that would have been observed if no treatment had been given. A major limitation of the RPSFT is the “common treatment effect” assumption, that is, the treatment effect received by switchers must be the same as the treatment effect received by patients initially randomised to the experimental group (i.e., similar efficacy of treatment whether initiated on or switched to at a later time when the disease prognosis for a patient might have changed).

- Two-stage method: Assumed that the trial is randomised up until the point of disease progression. Firstly, a treatment effect specific to switching patients is estimated and the survival times of these patients are adjusted, subsequently allowing the treatment effect specific to experimental group patients to be estimated. This approach makes use of structural nested failure time models (SNM) with g-estimation to estimate the treatment effect in switchers. Again, this approach has a similar limitation to the IPCW method where it assumes there is no “unmeasured cofounders”. Generally, the IPCW method is preferred to the two-stage method as no effort is made to adjust for any time-dependent confounding that occurs between disease progression and the time of switch.

All three adjustment methods were explored and are reported in Section B.2.6.4. The RPFST adjustment was deemed most appropriate and used in the base case.

B.2.5. Critical Appraisal of the relevant clinical effectiveness evidence

A quality assessment of the COSMIC-311 trial is summarised in Table 10. The COSMIC-311 trial was designed and undertaken according to the standards of good clinical practices, with adequate randomisation and blinding procedures. Please see Appendix D for a detailed quality assessment.

Table 10: Quality assessment results for the COSMIC-311 trial

Trial	The COSMIC-311 trial
Was randomisation carried out appropriately?	Yes
Was the concealment of treatment allocation adequate?	Yes
Were the groups similar at the outset in terms of prognostic factors?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes
Were there any unexpected imbalances in drop-outs between the groups?	No
Is there any evidence to suggest the authors measured more outcomes than they reported?	No (company-sponsored study)
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate measures used to account for missing data?	Yes/Yes/Yes

Abbreviations: ITT – Intention to treat

As patients were permitted to crossover to receive cabozantinib when disease progressed, there is potential bias of OS results due to this treatment switching. Therefore, as described in Section B.2.4.2.2, further analysis of the COSMIC-311 data was required to estimate the unbiased survival benefit of cabozantinib treatment compared to the placebo arm, adjusting for placebo patients crossing over to subsequent cabozantinib treatment. In Section B2.6.4 all results from crossover adjusted methods are reported.

Crossover in the trial also had an impact on the collection and interpretation of HRQoL in COSMIC-311. Patients that crossed over from the placebo arm to the treatment arm discontinued follow up of HRQoL. The primary criteria for treatment crossover were

radiographic progression of placebo patients, therefore many placebo patients did not have HRQoL data collected when they were in the post-progression state. There is no treatment crossover adjustment for HRQoL results and it could be the case that patients who remained in the placebo arm had a different prognosis than those who crossed over, which makes the interpretation of HRQoL difficult for placebo patients' post-progression. See Section B.3.4 for further description and analysis.

In addition to the crossover adjustments, other potential limitations including incomplete survival follow-up and high censoring in the placebo arm, may also impact the interpretability of the overall results for OS.

B.2.6. Clinical effectiveness results of the relevant studies

B.2.6.1 Primary Endpoint – Progression Free Survival

Table 11 presents a summary of the CCO1 (19th August 2020) and CCO2 (8th February 2021) analyses for PFS and ORR. The prespecified single interim analysis of PFS was planned at the time of the primary ORR analysis, although it was not expected that a mature PFS endpoint would be met at the time of this analysis. Thus, the prespecified interim analysis of the multiple primary endpoint PFS was conducted on the ITT population, at the time of the primary analysis of ORR in the overall response rate ITT at CCO1.

The study met the primary endpoint of PFS at the prespecified interim analysis in the ITT population. The HR, adjusted for stratification factors (per IxRS), was 0.22 (96% CI: 0.13, 0.36; stratified log-rank p-value <0.0001). The KM estimates for median duration of PFS were 11 months in the cabozantinib arm and 1.9 months for the placebo arm.

Given that the study demonstrated significant improvement in the primary endpoint of PFS at the prespecified interim analysis at CCO1, enrolment stopped, and the last subject was randomised on 2nd February 2021 with a total of 258 patients (170 patients in the cabozantinib arm and 88 patients in the placebo arm). Sites remained blinded through the primary efficacy analyses.

Subsequent efficacy analyses based on a second data cut-off date of 8th February 2021 (CCO2) were consistent with those of the corresponding primary analyses at CCO1 and showed that cabozantinib provides consistent and favourable outcomes across multiple efficacy analyses in this population of RAI-refractory DTC patients who had progressed after prior VEGFR-targeted therapy. At CCO2, a sustained improvement in PFS was demonstrated for cabozantinib versus placebo (11 months vs 1.9 months, HR=0.22, 96% CI 0.15, 0.32 p-value <0.0001) see Figure 5.

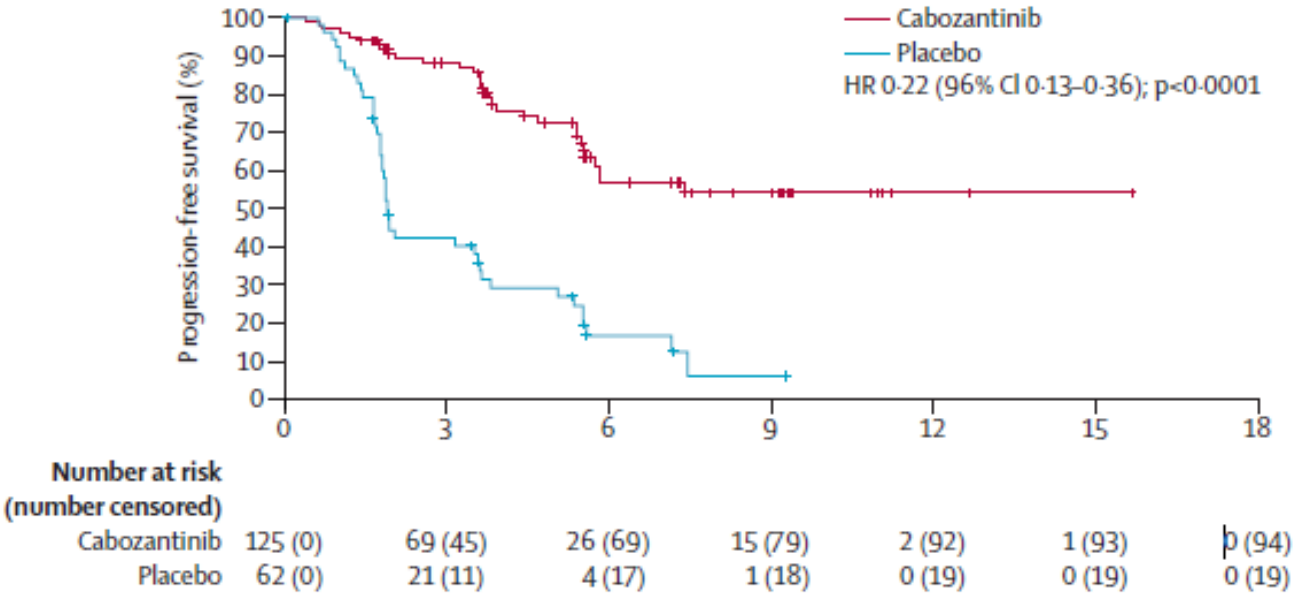
Table 11: Progression-free survival per BIRC (ITT population)

	CCO2* **(N = 258)**		**CCO1**** **(N = 187)**	
	Cabozantinib (N = 170)	**Placebo (N = 88)**	**Cabozantinib (N = 125)**	**Placebo (N = 62)**
Number (%) of subjects				
Censored	[redacted]	[redacted]	[redacted]	[redacted]
Receipt of local radiation to soft tissue for DTC	[redacted]	[redacted]	[redacted]	[redacted]
No post-baseline ATA[a]	[redacted]	[redacted]	[redacted]	[redacted]
No event by last ATA	[redacted]	[redacted]	[redacted]	[redacted]
2 or more missed ATA prior to event	[redacted]	[redacted]	[redacted]	[redacted]
Systemic NPACT	[redacted]	[redacted]	[redacted]	[redacted]
Event	62 (36)	69 (78)	31 (25)	43 (69)
Death	[redacted]	[redacted]	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]	[redacted]	[redacted]
Duration of PFS (months)				
Median (96% CI)	11.0 (7.4, 13.8)	1.9 (1.9, 3.7)	NE (5.7, NE)	1.9 (1.8, 3.6)
25th percentile, 75th percentile[b]	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Observed p-value (stratified log-rank test)[c]	<0.0001		<0.0001	
Hazard ratio (96% CI; stratified)[c,d]	[redacted]		[redacted]	
Hazard ratio (96% CI; stratified)[c,d]	0.22 (0.15, 0.32)		0.22 (0.13, 0.36)	
Observed p-value (unstratified log-rank test)	[redacted]		[redacted]	
Hazard ratio (95% CI; unstratified)[d]	[redacted]		[redacted]	
Hazard ratio (96% CI; unstratified)[d]	[redacted]		[redacted]	

KM landmark estimates (% of subjects event-free) at:				
3 months	[redacted]	[redacted]	[redacted]	[redacted]
6 months	[redacted]	[redacted]	56.9	16.9
9 months	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]

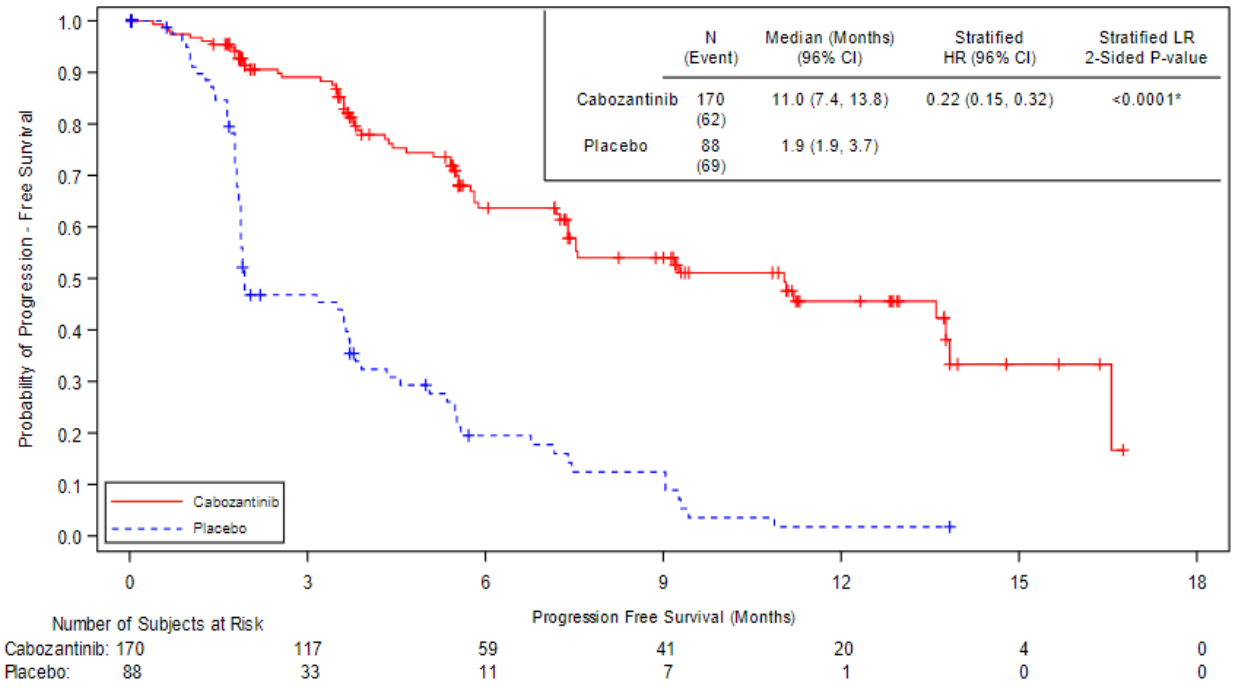
* 8th February 2021 cut-off
** 19th August 2020 cut-off
Abbreviations: ATA – Adequate tumor assessment; BIRC – Blinded independent radiology committee; CI – Confidence interval; DTC – differentiated thyroid cancer; HR – Hazard ratio; ITT – Intent-to-treat; IxRS – Interactive voice/web response system; KM – Kaplan-Meier; NPACT – Nonprotocol anticancer therapy; ORR – Objective response rate; PD – Disease progression; PFS – Progression-free survival
+ indicates a censored observation (please see PFS censoring rules in XL184-311 CSR, Section 9.7.1.2.2)
a. In the Full ITT population, 11 cabozantinib and 8 placebo subjects were enrolled too close to the data cut cutoff date to have had a post-baseline tumour assessment. Four cabozantinib subjects decided to withdraw from treatment before any postbaseline tumor assessment. In addition, 3 subjects in the cabozantinib arm (1807-3002, 3808-3111, and 3907-3338) and 1 subject in the placebo arm (3905-3275) died before their first post-baseline scan.
b. Percentiles were based on KM estimates.
c. Stratification factors (per IxRS) comprise receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years).
d. Estimated using the Cox proportional-hazard model (adjusted for stratification factors if applicable). HR < 1 indicated PFS in favor of cabozantinib.
Source: XL184-311 CSR (30th April 2020)[46] and XL184-311 CSR Addendum 1 (21st May 2021)[51] and Brose et al, 2021[48]

Figure 4: Kaplan-Meier plot of progression-free survival per BIRC through 19 August 2020 (ITT Population) (CCO1; N=187)



Abbreviations: BIRC – Blinded independent radiology committee; CI – Confidence interval; HR – Hazard ratio; ITT – Intent-to-treat; IxRS – Interactive voice/web response system; LR, – Log-rank test.
Stratification factors (per IxRS) comprise receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years).
Source: XL184-311 CSR (30th April 2020)[46] and Brose et al, 2021[48]

Figure 5: Kaplan-Meier plot of progression-free survival per BICR through 08 February 2021 (Full ITT Population) (CCO2; N=258)



Abbreviations: BIRC – Blinded independent radiology committee; CI – Confidence interval; HR – Hazard ratio; ITT – Intent-to-treat; IxRS – Interactive voice/web response system; LR – Log-rank test.
+ indicates value from censored observation
Stratification factors (per IxRS) comprise receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years).
Source: XL184-311 CSR Addendum 1 (21st May 2021)[51]

B.2.6.2 Primary Endpoint – Objective Response Rate (ORR)

Analysis of the primary endpoint ORR per RECIST 1.1 as determined by BIRC were performed on the entire ITT population and also in the first interim analysis. ORR results are presented in Table 12 for CCO1 and CCO2.

In CCO2, the ORR was 11% (95% CI: 6.9, 16.9) and 0% (95% CI: 0.0, 4.1) for subjects in the cabozantinib and placebo arms, respectfully. Of the reported objective responses, 18 out of 19 in the cabozantinib arm were partial responses (PRs) with there being 1 complete response. There was a higher rate of stable disease (SD) in the cabozantinib arm relative to the placebo arm (68.8% vs 38.6%, respectively). The frequency of progressed disease as best response was lower in the cabozantinib arm compared with the placebo arm (6.5% vs 47.7%, respectively), indicating a low incidence of primary refractory disease to cabozantinib treatment in this study

population. The KM estimate of median (range) duration of objective response (DOR) per BIRC was 10.2 (9.3 TO NE) months in the cabozantinib arm. The median (range) time from randomisation to the first objective response per BIRC was 3.6 (1.74, 7.52) months in the cabozantinib arm.[51]

The best percentage change from baseline in tumour target lesion size (investigator-determined according to RECIST 1.1) for CCO1 is depicted in Figure 6 for both cabozantinib and placebo and CCO2 is shown in Figure 7. In CCO1, among subjects in the OITT population with at least one baseline and at least one post-baseline target lesion assessment, 44/58 (76%) in the cabozantinib arm and 9/31 (29%) in the placebo arm had a postbaseline reduction in the sum of target lesion diameters (SoD). In CCO2, among subjects in the Full ITT population with at least one baseline and at least one post-baseline target lesion assessment, 115/144 (80%) in the cabozantinib arm and 18/76 (24%) in the placebo arm had a postbaseline reduction in the sum of target lesion diameters (SoD).

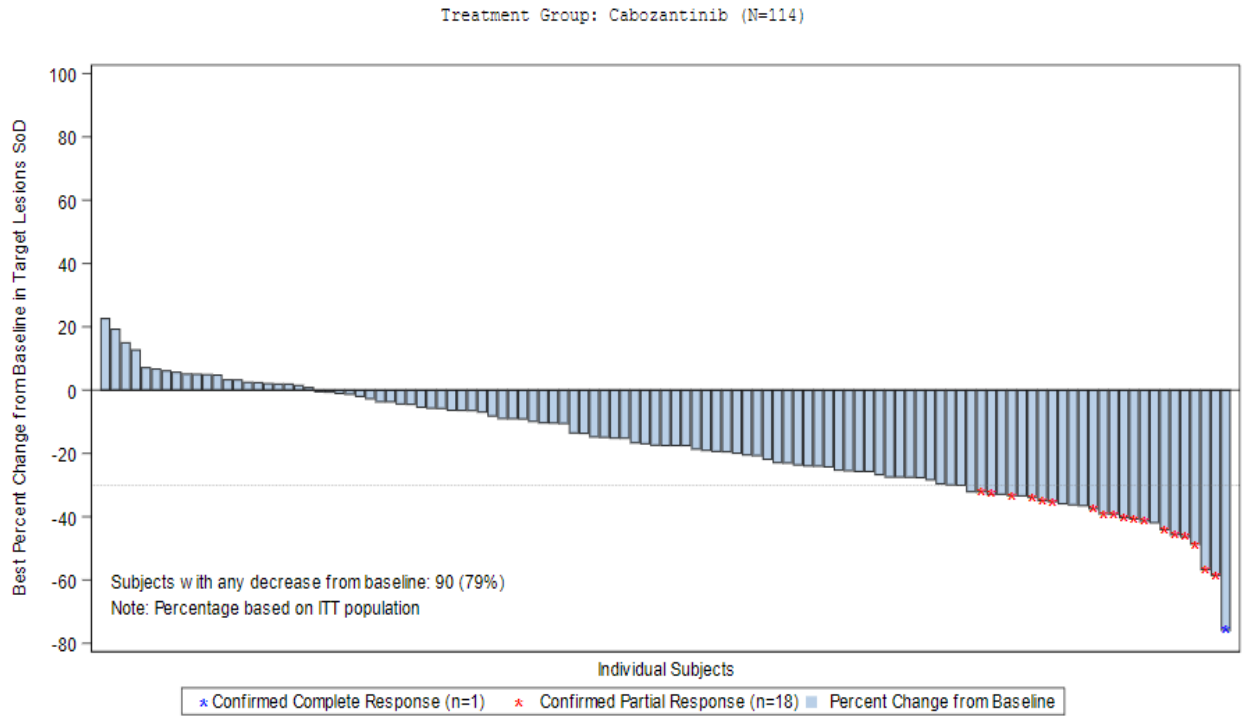
Table 12: Objective response rate per BIRC (ITT population)

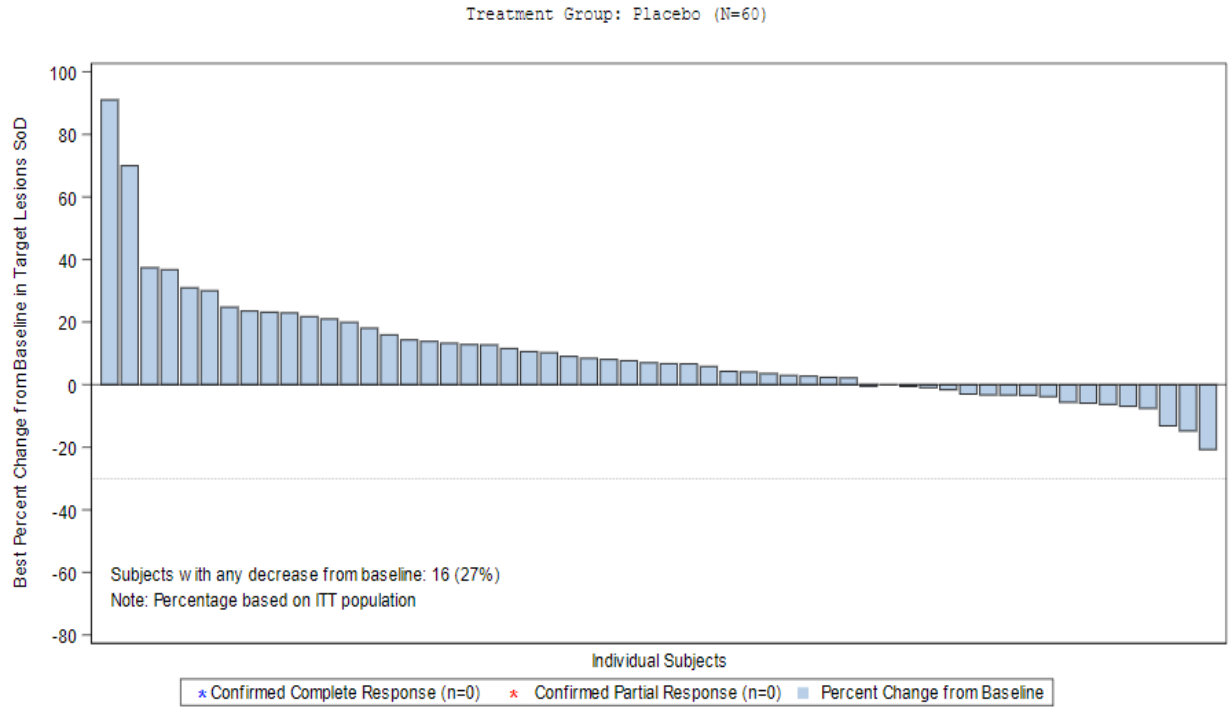
	CC02 (N = 258)		**CCO1 (N = 187)**	
	Cabo (N = 170)	**Placebo (N = 88)**	**Cabo (N = 125)**	**Placebo (N = 62)**
Best overall response, n (%)[a]				
Confirmed complete response (CR)	1 (0.6)	0	0	0
Confirmed partial response (PR)	18 (10.6)	0	11 (9%)	0
Stable disease (SD)	117 (68.8)	34 (38.6)	76 (61)	21 (34)
Progressive disease (PD)	11 (6.5)	42 (47.7)	8 (6)	31 (50)
No disease (NA)	1 (0.6)	0	1 (1)	0
Unable to evaluate (UE)	3 (1.8)	1 (1.1)	2 (2)	1 (2)
Missing	19 (11.2)	11 (12.5)	27 (22)	9 (15)
Objective response rate (CR+PR), n (%)	19 (11)	0	11 (9)	0
95% CI	6.9, 16.9	0.0, 4.1	4.5, 15.2	0.0, 5.8
Observed unstratified Fisher exact test p-value	0.0003		0.017	
Disease stabilisation rate (ORR+SD ≥ 16 weeks), n (%)	90 (52.9)	17 (19.3)	54 (43)	10 (16)
95% CI	[redacted]	[redacted]	34.4-52.5	8.0-27.7

Duration of Objective Response per BIRC (KM), median (range), months	10.2 (9.3, NE)	NA	NR	NA
Time to Objective Response per BIRC, median (range) time from randomisation, months[b]	3.581 (1.74, 7.52)	NA	1.9 (1.8-3.6)	NA

Abbreviations**:** BIRC – Blinded Independent Radiology Committee; CI – Confidence interval; CMH – Cochran Mantel-Haenszel; OITT – Overall response rate intent-to-treat; IxRS – Interactive voice/web response system; RECIST – Response Evaluation Criteria in Solid Tumors; NR – Not reached; NA – Not applicable.
a) Best overall response was assessed based on RECIST 1.1 criteria and was calculated based on subjects in the OITT population. Note that a CR or PR was not considered as an objective response if a subject progressed or received subsequent anticancer therapy prior to the first CR or PR. To be classified as a CR or PR, confirmation of response must have occurred > 28 days after the response was first observed.
b) Time to objective response is an arithmetic summary amongst those with an objective response and is defined as time from randomization to the first CR or PR that is subsequently confirmed.
Source: XL184-311 CSR Addendum 1 (21st May 2021)[51] and Brose et al, 2021[48]

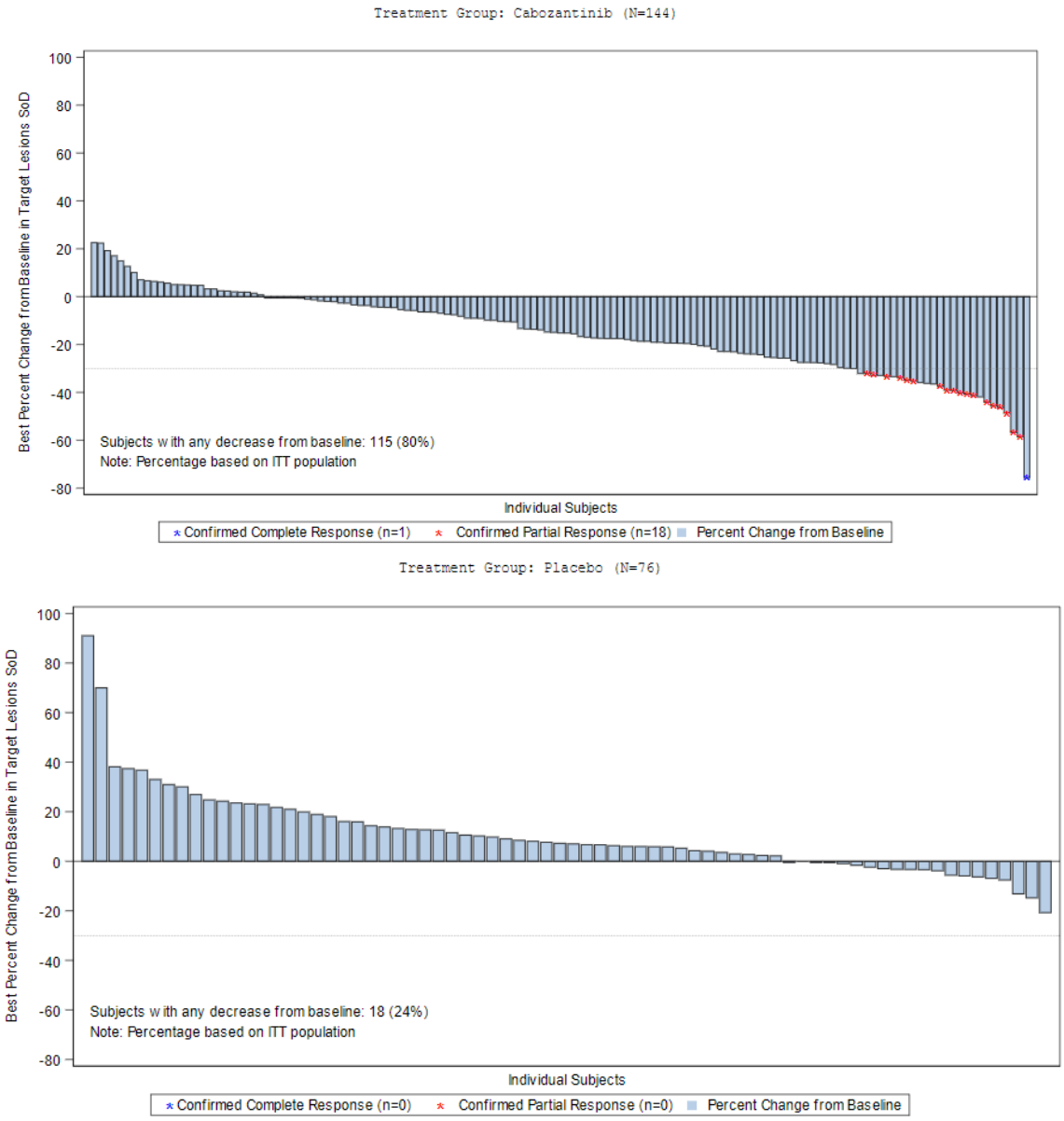
Figure 6: CCO1 Waterfall plot of best percentage change in tumour target lesion size from baseline per BIRC*





*Subjects with at least one baseline and at least one post baseline target lesion assessment.
Source: XL184-311 CSR (30th April 2020)[46] and Brose et al, 2021[48]

Figure 7: CCO2 Waterfall plot of best percentage change in tumour target lesion size from baseline per BIRC*



*Subjects with at least one baseline and at least one post baseline target lesion assessment.
Source: XL184-311 CSR Addendum 1 (21st May 2021)[51]

B.2.6.3 Secondary Endpoint – Overall Survival (OS)

OS was a secondary endpoint in this trial. The analysis of OS was descriptive and non-inferential as OS was not a controlled endpoint for the study. Given the potential for crossover from placebo to cabozantinib and the potential for receipt of subsequent non-protocol anticancer therapy (NPACT), it was acknowledged that the interpretation

of the OS results is limited. Therefore, OS was only descriptive and not a controlled endpoint for the study. The primary purpose of the OS analyses was to evaluate the potential for detriment to survival with cabozantinib treatment.

Survival status as of CCO1 was determined for all 187 randomised subjects. However, interpretation of this analysis was limited since only 17 subjects (14%) in the cabozantinib arm and 14 subjects (23%) in the placebo arm had events. The remaining 108 subjects (86%) in the cabozantinib arm and 48 subjects (77%) in the placebo arm were censored at their last known alive date including the 6 cabozantinib and 3 placebo subjects who died after the clinical cut-off date. A total of 2 subjects withdrew full consent including for survival follow-up. The placebo subjects who crossed over to receive cabozantinib were analysed under the placebo arm.

The median time of follow-up through 19 August 2020 (CCO1) was 6.24 months in the ITT population. No values were imputed. Of note, the placebo arm included 19 subjects who subsequently crossed over to receive cabozantinib; these subjects were not censored at the time of crossover and were analysed under the randomised placebo arm for OS analysis under intent-to-treat principles.

The analysis demonstrated a trend for longer OS for subjects in the cabozantinib arm in CCO1 compared with the placebo arm (Table 13): the HR, adjusted for stratification factors (per IxRS), was 0.54 (95% CI: 0.27, 1.11). The KM estimates for median duration of OS were not estimable in either arm (Table 13). The proportion of subjects alive at 6 months was 84.8% in the cabozantinib arm compared with 73.4% in the placebo arm.

At the time of CCO2, OS was immature. Interpretation of the OS data is limited due to the low number of events: 58 deaths at CCO2 (37 cabozantinib, 21 placebo), 133 patients (78%) in the cabozantinib arm and 67 patients (76%) in the placebo arm were censored at their last known alive dates. However, the Full ITT population analysis demonstrated a trend for longer OS for patients in the cabozantinib arm compared with the placebo arm. It should be noted that the p-values provided should not be used for statistical inferences as OS was not a controlled endpoint. The Kaplan-Meier curves in Figure 9 illustrate OS for the full ITT population. The analysis demonstrated

a trend for longer OS for patients in the cabozantinib arm compared with the placebo arm (HR = 0.76, 95% CI 0.45, 1.31). The proportion of patients alive at 6 months was [redacted]% in the cabozantinib arm compared with [redacted]% in the placebo arm. The placebo arm included 40 patients (45%) who crossed over to receive cabozantinib, eight of whom had an event. The placebo crossover patients were not censored at the time of crossover and were analysed under the randomised placebo arm for OS analysis under ITT principles. For these 40 patients who crossed over to receive cabozantinib upon BIRC-confirmed radiographic progression, selected demographic and baseline characteristics were re-established immediately prior to crossover. Analyses were conducted on the OS data to estimate the unbiased survival benefit of cabozantinib treatment compared to the placebo, as previously mentioned (Section B.2.4.2.2), adjusting for placebo patients crossing over to subsequent cabozantinib treatment. These results can be found in Section B.2.6.4.

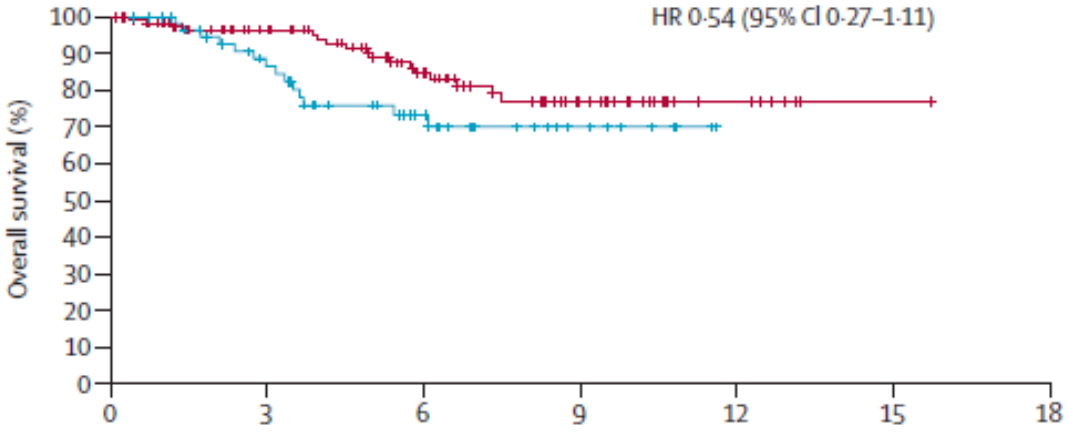
Table 13: Overall survival CCO1 and CCO2 ITT population

	CCO2*		**CCO1****	
	Cabozantinib (N=170)	**Placebo (N=88)**	**Cabozantinib (N=125)**	**Placebo (N=62)**
Number of subjects (%)				
Censored	[redacted]	[redacted]	[redacted]	[redacted]
Alive	131 (77)	67 (76)	NR	NR
Death after data cut-off date	[redacted]	[redacted]	NR	NR
Death	37 (22)	21 (24)	17 (14)	14 (23)
Duration of overall survival (months)[a]				
Median (95% CI)	19.4 (15.9, NE)	NE (NE, NE)	NE (NE, NE)	NE (NE, NE)
25th percentile	[redacted]	[redacted]	[redacted]	[redacted]
75th percentile	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Observed p-value (stratified log-rank test)[b]	[redacted]		[redacted]	
Hazard ratio (95% CI; stratified)[b,c]	0.76 (0.45, 1.31)		0.54 (0.27, 1.11)	
Observed p-value (unstratified log-rank test)	[redacted]		[redacted]	
Hazard ratio (95% CI; unstratified)[c]	[redacted]		[redacted]	

KM landmark estimates (% of subjects event-free) at:				
3 months	[redacted]	[redacted]	[redacted]	[redacted]
6 months	[redacted]	[redacted]	84.8	73.4
9 months	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]

* 8th February 2021 cut-off
** 19th August 2020 cut-off
Abbreviations: CC01- Clinical cut-off 1; CC02-Clinical cut-off 2; CI – Confidence; HR – Hazard ratio, ITT – Intent-to-treat; LR – Log-rank test, NE – Not estimable; NR – Not reported; OS – Overall survival.
+indicates a censored observation (please see OS censoring rules in XL184-311 CSR, Section 9.7.1.4.1).
a Percentiles were based on K-M estimates.
b Stratification factors based on IxRS were receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years). c Estimated using the Cox proportional-hazard model (adjusted for stratification factors if applicable). HR < 1 indicated OS in favour of cabozantinib.
d In the Full ITT population and Primary Analysis subset, maximum duration of OS in the placebo arm was 17.28 months at CCO2.
Source: XL184-311 CSR (30th April 2020)[46] and XL184-311 CSR Addendum 1 (21st May 2021)[51] and Brose et al, 2021[48]

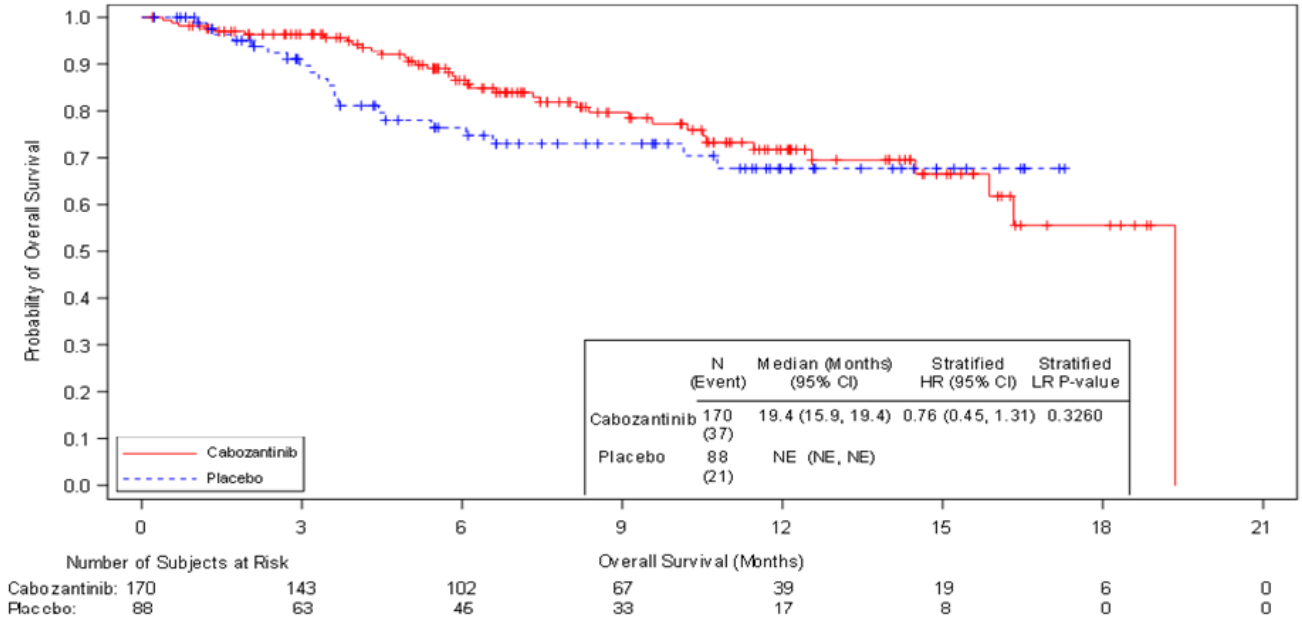
Figure 8: Kaplan-Meier plot of OS (CCO1; N=187)



Number at risk (number censored)							
Cabozantinib	125 (0)	90 (31)	54 (58)	24 (84)	7 (101)	1 (107)	0 (108)
Placebo	62 (0)	42 (13)	24 (25)	9 (39)	0 (48)	0 (48)	0 (48)

Abbreviations: CI – Confidence; HR – Hazard ratio, ITT – Intent-to-treat; LR – Log-rank test, NE – Not estimable; OS – Overall survival
+ Indicates censored observation
The upper limit of the 95% CI for median OS should be interpreted as NE.
The last remaining subject in the cabozantinib arm had an event leading the survival probability to 0% as no subject remained at risk anymore.
Source: XL184-311 CSR Addendum 1 (21st May 2021)[51] and Brose et al, 2021[48]

Figure 9: Kaplan-Meier plot of OS (CCO2; N=258)



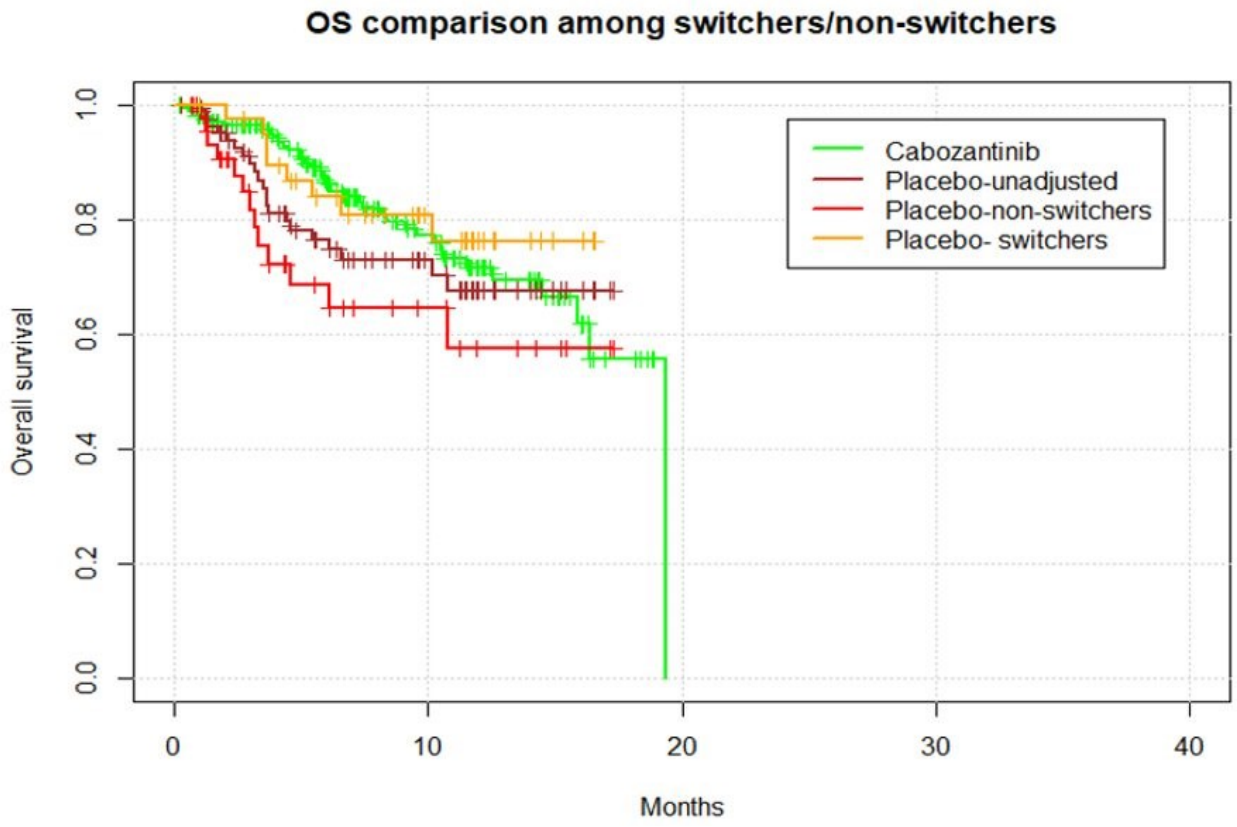
Abbreviations: CI – Confidence; HR – Hazard ratio, ITT – Intent-to-treat; LR – Log-rank test, NE – Not estimable; OS – Overall survival
+ Indicates censored observation
The upper limit of the 95% CI for median OS should be interpreted as NE.
The last remaining subject in the cabozantinib arm had an event leading the survival probability to 0% as no subject remained at risk anymore.
Source: XL184-311 CSR Addendum 1 (21st May 2021)[51]

B.2.6.4 Crossover Adjustment Analyses

As described in section B.2.4.2.2, further analysis of the COSMIC-311 data was required to estimate the unbiased survival benefit of cabozantinib treatment compared to the placebo arm, adjusting for placebo patients crossing over to subsequent cabozantinib treatment.

An ad-hoc sensitivity analysis looking at OS results for placebo-unadjusted for crossover, placebo treatment switchers and placebo non-switchers was conducted. Figure 10 outlines the KM results for these sub-groups in comparison to patients in the cabozantinib only arm. It is clear that patients who did not switch treatment were associated with significantly worse OS, with a stratified HR of 0.42 (95% CI 0.22, 0.81), although caution is to be taken in interpreting these results. There is likely to be bias as placebo non-switchers may have had a differing prognosis at baseline than treatment switchers. Thus a crossover adjustment was conducted in the base case.

Figure 10: Overall survival Kaplan-Meier curves of switched patients and non-switched patients



Overall, the treatment crossover adjusted analyses indicated a trend towards improved survival in the COSMIC-311 data available. However, it should be noted that there was incomplete survival follow-up, as 77% of the total sample size was censored. All three adjustments methods gave similar results (see Table 14 and Figure 11 below), although the RPSFT and the two-stage results are more likely most appropriate for adjusting treatment crossover in the COSMIC-311 trial.[54] The relative efficacy of cabozantinib vs placebo before adjustments was estimated with a HR of 0.76 (95% CI; 0.45, 1.31 / Full ITT population). The HR for cabozantinib vs placebo-RPSFT adjusted was 0.65 (95% CI; 0.38, 1.13), and 0.7 (95% CI; 0.41, 1.22) for cabozantinib vs placebo-two-stage adjusted, and 0.68 (95% CI; 0.37, 1.27) for cabozantinib vs placebo-IPCW adjusted. The mean difference (MD) in OS between cabozantinib and placebo-RPSFT adjusted was estimated as 10.19 (95% CI; -6.95, 27.33) months, and 8.33 (95% CI; -8.81, 25.47) months for cabozantinib vs placebo-

two-stage adjusted and 5.82 (95% CI; -11.7, 23.34) months for cabozantinib vs placebo-IPCW adjusted.

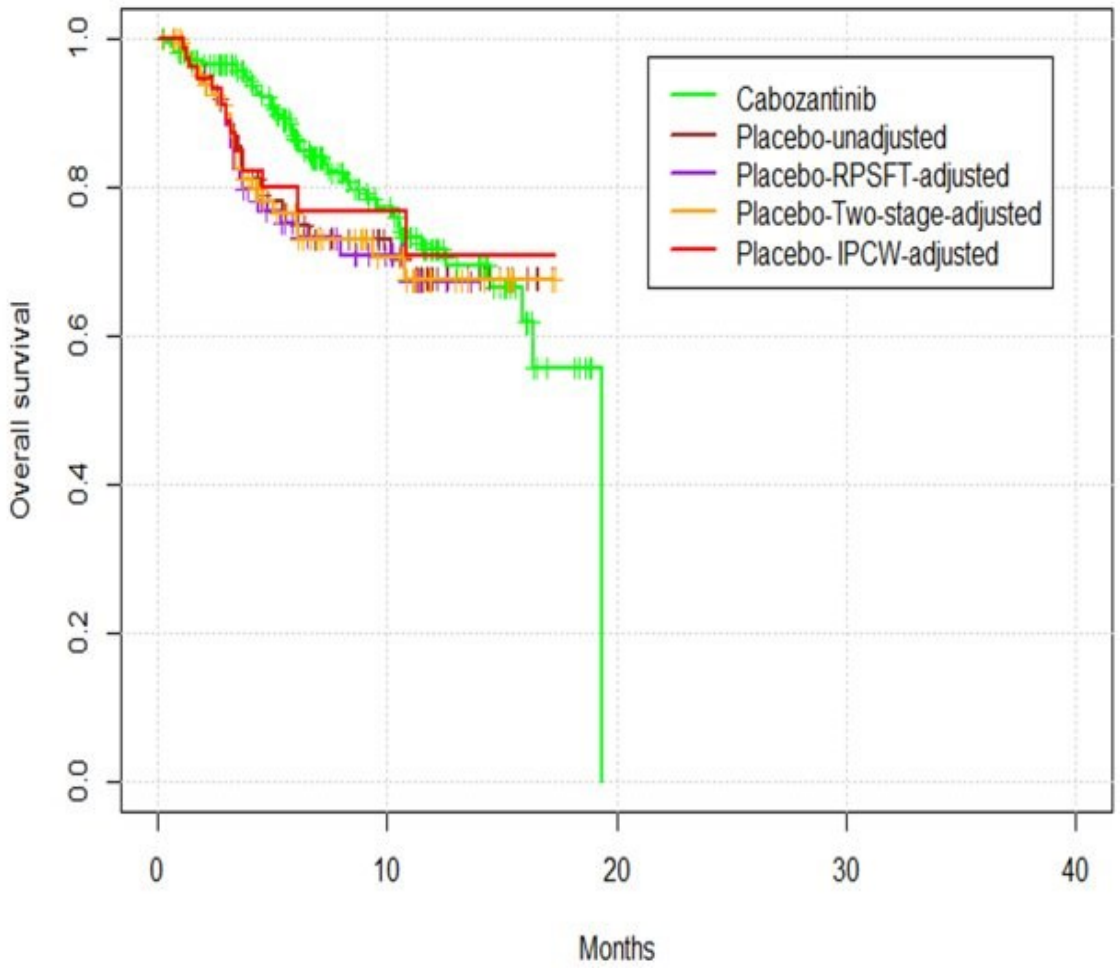
Although the RPSFT and the two-stage results are more likely most appropriate for adjusting treatment crossover in the COSMIC-311 trial, the RPSFT method has been used as the base case because it was in line with previous NICE submissions, in particular TA535.[55]

Table 14: Overall survival (95% CI) estimates for cabozantinib vs placebo before and after adjustments

Distribution	Placebo-unadjusted (95% CI)	Placebo-RPSFT (95% CI)	Placebo-two-stage (95% CI)	Placebo-IPCW (95% CI)
Stratified HR (naïve 95% CI)	0.76 (0.45, 1.31)	0.65 (0.28, 1.53)	0.70 (0.41, 1.22)	0.68 (0.37, 1.27)
Mean survival – Cabozantinib	37.58 (27.08, 50.74)	37.58 (27.08, 50.74)	37.58 (27.08, 50.74)	37.58 (27.08, 50.74)
Mean survival – Placebo	30.45 (20.89, 45.71)	27.39 (18.38, 41.15)	29.25 (18.83, 43.47)	31.76 (19.5, 51.59)
Mean difference - cabozantinib vs. Placebo	7.13 (-10.01, 24.27)	10.19 (-6.95, 27.33)	8.33 (-8.81, 25.47)	5.82 (-11.7, 23.34)

Abbreviations: CI – Confidence Interval; IPCW – Inverse Probability of Censoring Weights; RPSFT – Rank Preserved Structural Failure Time; MDs – Mean Difference

Figure 11: Overall survival Kaplan-Meier curves of the 3 methods of crossover adjustment



Abbreviations: RPSFT – Rank Preserved Structural Failure Time; IPCW – Inverse Probability of Censoring Weights

B.2.6.5 Health Related Quality of Life

HRQoL was measured throughout COSMIC-311 using the EuroQoL-5-dimension with 5 levels (EQ-5D-5L) instrument. Patients completed the questionnaires at baseline (before receiving the treatment or control), and post-baseline assessments were collected every 4 weeks until week 25 and then every 8 weeks thereafter. EQ-5D-5L questionnaires were discontinued post progression and for patients who transitioned to the crossover phase. The HRQoL results are only reported for CCO1.

The EQ-5D-5L questionnaire completion rate at baseline was 98% in the cabozantinib arm and 100% in the placebo arm and remained above 80% in each treatment arm through Week 33. Beyond Week 33 there were fewer than 5 patients in the placebo arm. The effect size for change from baseline was calculated as mean of change in score/pooled SD for baseline scores. An effect size ≥0.3 was considered potentially clinically meaningful, meaning that an EQ-Index value difference from baseline above 0.3 is considered clinically meaningful. The minimal important difference (MID) threshold is 0.3, which is consistent with previous HRQoL analyses.[56]

On all dimensions of the EQ-5D-5L, changes from baseline in patients in the cabozantinib and in the placebo arms did not show any statistically or clinically meaningful treatment difference. This is not consistent with other oncology treatments as usually an early deterioration in QoL is experienced by patients due to AE's. At baseline, mean EQ-VAS scores for the ITT population were [redacted] in the cabozantinib arm and [redacted] in the placebo arm respectively. Subsequently, QoL remained stable throughout the duration of the treatment, up to time points with less than 5 patients by arm.

Figure 12 shows the mean change of EQ-5D index scores from baseline. At baseline, mean EQ-Index scores were [redacted] in the cabozantinib arm and [redacted] in the placebo arm. All treatment differences in mean change from baseline EQ-Index values were less than [redacted] through Week 33, which is below the MID threshold. Results post week 33 should not be interpreted due to the very low sample size in the placebo arm.

Figure 12: Mean (SE) change from baseline of EQ-Index score (CCO1 ITT population)

Abbreviations**:** ITT – Intent–to–treat; post–BL – Post–baseline; SE – Standard error; W – Week
Source: Global Value Dossier, Ipsen 2022. Median follow up: CCO1 ITT – 6.24 months.
Source: XL184-311 CSR (30th April 2020)[46]

Figure 13 outlines the mean change from baseline of EQ-VAS score in the ITT population. At baseline, mean EQ-VAS scores were [redacted] in the cabozantinib arm and [redacted] in the placebo arm. All treatment differences in mean change from baseline EQ-VAS values were less than 7 through Week 33, again below the MID threshold.

Figure 13: Mean (SE) change from baseline of EQ-VAS score (CCO1 ITT population)

Abbreviations**:** ITT – Intent–to–treat; LSMean – Least squares means; SD – Standard deviation; SE – Standard error; VAS – Visual analogue scale.
Source: XL184–311 CSR (30th April 2020)[46]

Table 15 shows a summary of the HRQoL results from COSMIC-311. There was no potential clinically meaningful HRQoL difference between cabozantinib and placebo post week 33. After week 33, EQ-5D-5L questionnaires were collected in less than 5 patients in the placebo arm. Therefore, it is difficult to interpret results post week 33. Overall, the treatment of RAI refractory with cabozantinib has not shown a quality-of-life deterioration.

Table 15: EQ-VAS and EQ-Index scores: change from baseline, repeated measures analysis (CCO1 ITT population)

	EQ-5D Index	EQ-VAS
Cabozantinib n (N = 125)		
Cabozantinib least square means (SE)		
Placebo n (N = 62)		
Placebo least square means (SE)		
Difference in mean change		
Pooled SD		
P-value		
Effect size		

Abbreviations**:** EQ–5D – EuroQol five–dimension; EQ–VAS – EuroQol visual analogue scale; SE – Standard error. Source**:** XL184–311 CSR (30th April 2020)[46]

.

B.2.6.6 Sensitivity Analyses

Table 16 provides point estimates and 95% CIs of stratified HRs for the prespecified sensitivity analyses of PFS as described above in Section 2.4.2, in which additional or alternative clinical outcomes were considered to be events and various definitions of disease progression.

Table 16: PFS sensitivity analysis results

	No. of events/subjects (%) Median duration (mo)		Stratified Hazard Ratio	95% CI	Stratified Log-rank p-value
PFS Analysis	**Cabozantinib**	**Placebo**			
CCO1 (n=258)					
PFS-EP-1a	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EP-2b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EP-4c	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EA2-1d	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EA2-2e	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CCO2 (n=187)					
PFS-EP-1[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EP-2[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EP-4[c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EA2-1[d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS-EA2-2[e]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ATA – Adequate tumor assessments; CCO2 – Clinical Cutoff 2; CI – Confidence interval; ITT – Intent–to–treat; No – Number; NPACT – Non–protocol anticancer therapy; PFS – Progression–free survival
a PFS-EP-1: PFS analysis as of CCO2 (data cutoff of 08 February 2021).
b PFS-EP-2: Date of radiographic progression was based on the date of the scheduled visit, rather than the date of recorded progression.
c PFS-EP-4: Rather than being censored, subjects who experienced ≥ 2 consecutive missing scheduled ATA immediately prior to documented radiographic progression were classified as having an event at the date of the last ATA prior to the missing visits.
d PFS-EA2-1: Receipt of systemic NPACT was changed to "composite," resulting in an endpoint that comprised radiographic progression, death, or initiation of systemic NPACT (XL184-311 CSR, Section 9.7.1.2.2.2).
e PFS-EA2-2: Sensitivity analysis of PFS-EA2-1 similar to PFS-EP-4 (footnote "c" above).
Source: XL184-311 CSR Addendum 1 (21st May 2021)[51]

B.2.7. Subgroup Analysis

The subgroup analyses of PFS in CCO1 and CCO2 showed a consistently favourable effect of cabozantinib compared with placebo. At the point estimate, in CCO1 and CCO2 cabozantinib also showed a consistent OS benefit compared to placebo. Although, given the immaturity of the OS data confidence intervals are wide.

Sub-group analysis was performed for ORR in CCO1. Although, due to small sample size of the sub-groups, results are difficult to interpret. There were no objective responses reported in the placebo arm, subgroup analyses of ORR were not performed at CCO2.

Outcome data for the following subgroups has been presented in this section:

- Receipt of prior lenvatinib (Yes, No)
- Receipt of prior sorafenib (Yes, No)
- Receipt of prior sorafenib and lenvatinib (Yes, No)
- Histology (Papillary, Follicular)
- Bone, Important Visceral, Liver, Lung Metastases per Investigator (Yes, No)
- Race
- Gender

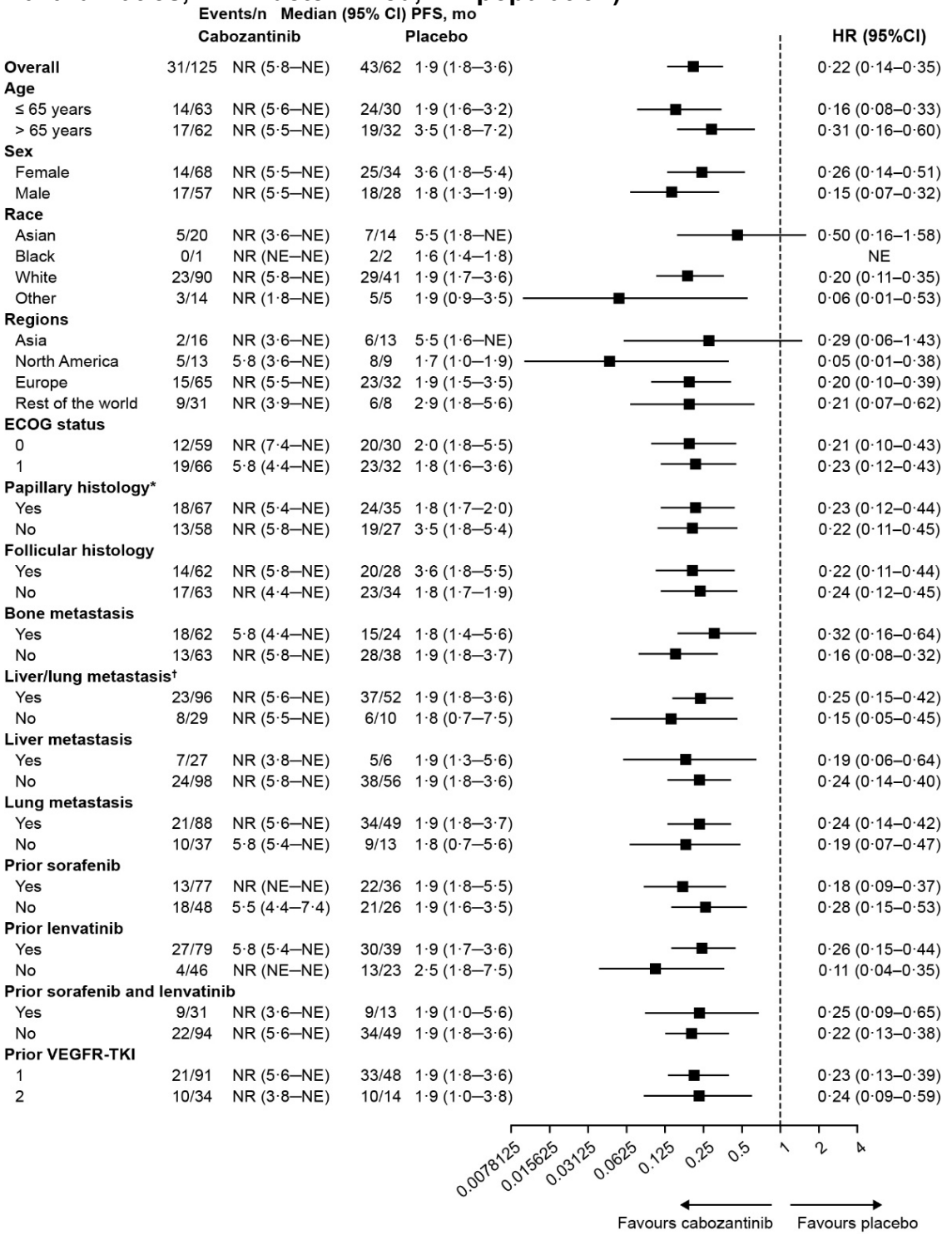
B.2.7.1 Progression Free Survival

Subgroup analyses is presented for PFS for both CCO1 and CCO2 in forest plots (Figure 14 and Figure 15). The PFS benefit was maintained across predefined subgroups with reasonable sample sizes. Almost all of the estimable HRs were below one and almost all upper limits of 95% CIs were also less than 1 (exceptions occurred in subgroups with a low number of patients). Of note, the favourable effect on PFS occurred regardless of the two stratification factors: receipt of prior lenvatinib (yes vs. no) or age at informed consent (≤65 years vs. >65 years).

The HR in CCO1 was 0.26 (95% CI: 0.15, 0.44) for subjects who received prior lenvatinib and 0.11 (95% CI: 0.04, 0.35) for those who did not receive prior lenvatinib. For age at informed consent, the HR was 0.16 (95% CI: 0.08, 0.33) for subjects ≤ 65 years and 0.31 (95% CI: 0.16, 0.60) for those > 65 years. The HR was similar for those patients who had received one or two prior VEGFR-TKI therapies i.e. HR 0.23 (95% CI:0.13-0.39) and HR 0.24 (95% CI: 0.09-0.58) respectively.

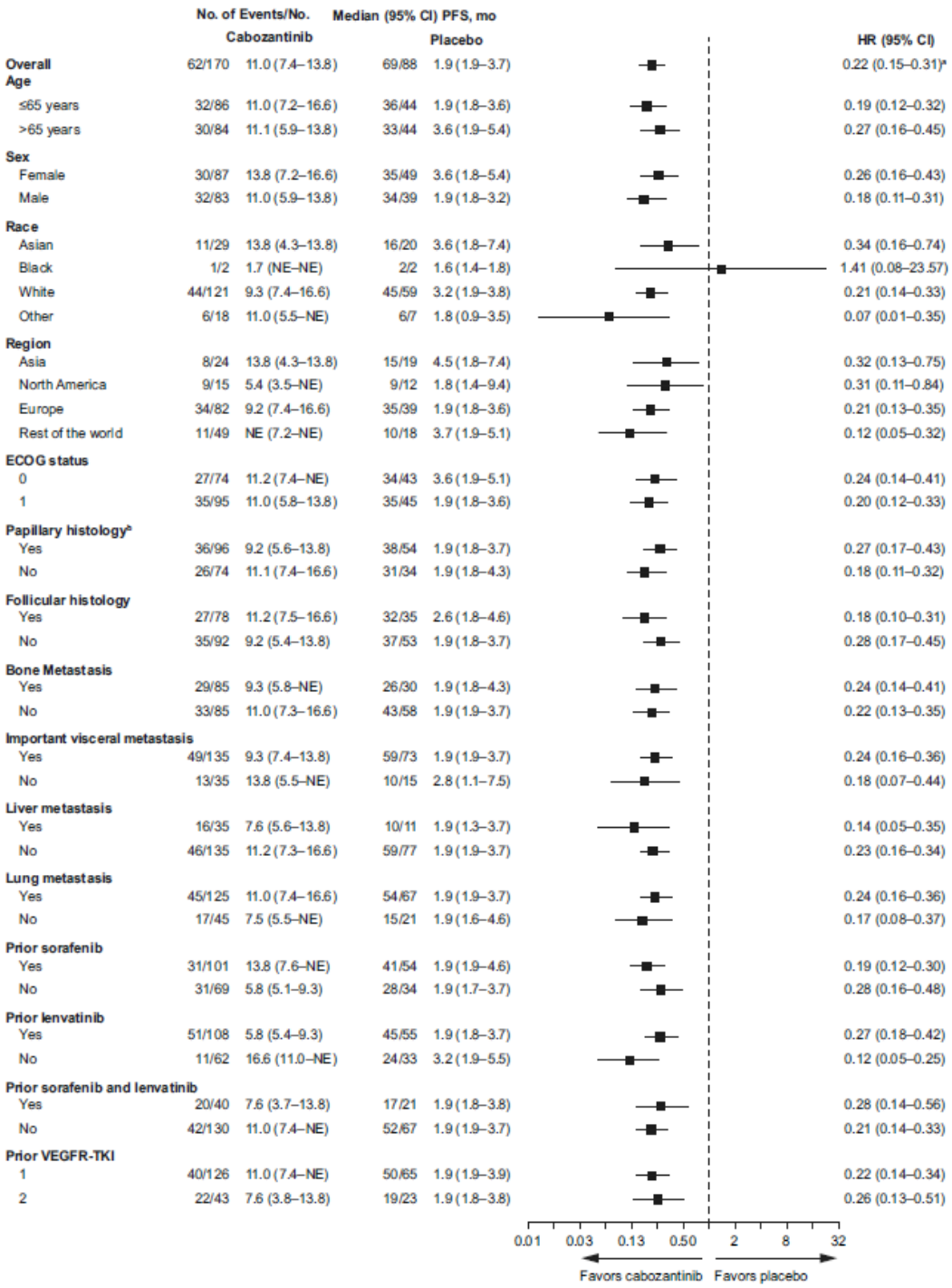
The HR for PFS in CCO2 was 0.27 (95%CI: 0.18, 0.42) for subjects who received prior lenvatinib and 0.12 (95%CI: 0.05, 0.25) for those who did not receive prior lenvatinib, indicating that cabozantinib can provide PFS benefit in both second- and subsequent-line RAI-refractory DTC. For age at informed consent, the HR was 0.19 (95%CI: 0.12, 0.32) for patients <65 years and 0.27 (95%CI: 0.16, 0.45) for those > 65 years.

Figure 14: CCO1 - Forest plots of subgroup analyses for PFS (Unstratified Hazard Ratios, BIRC-determined, ITT population)



Disease progression was assessed with the use of Response Criteria in Solid Tumours (RECIST), version 1·1 by blinded independent radiology committee. Hazard ratios are estimates from the Cox proportional hazards model and are unstratified with the exception of those for the overall population, which use the randomisation stratification factors. *17 patients were with papillary DTC had a follicular variant. †Important visceral metastasis. CI=confidence interval. ECOG=Eastern Cooperative Oncology Group. HR=hazard ratio. NE=not estimable. NR=not reached. TKI=tyrosine kinase inhibitor. VEGFR=vascular endothelial growth factor receptor.
Source: Brose et al, 2021[48]

Figure 15: CCO2 - Forest plots of subgroup analyses for PFS (Stratified Hazard Ratios, BIRC-determined, Full ITT population)



PFS in prespecified subgroups. aStratified hazard ratio. bThirty-two patients with papillary differentiated thyroid cancer had a follicular variant.
Abbreviations: CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; HR, hazard ratio; NE, not estimable; No.,
number; PFS, progression-free survival; TKI, tyrosine kinase inhibitor; VEGFR, vascular endothelial growth factor receptor
Source: Brose et al. 2022[57,58]

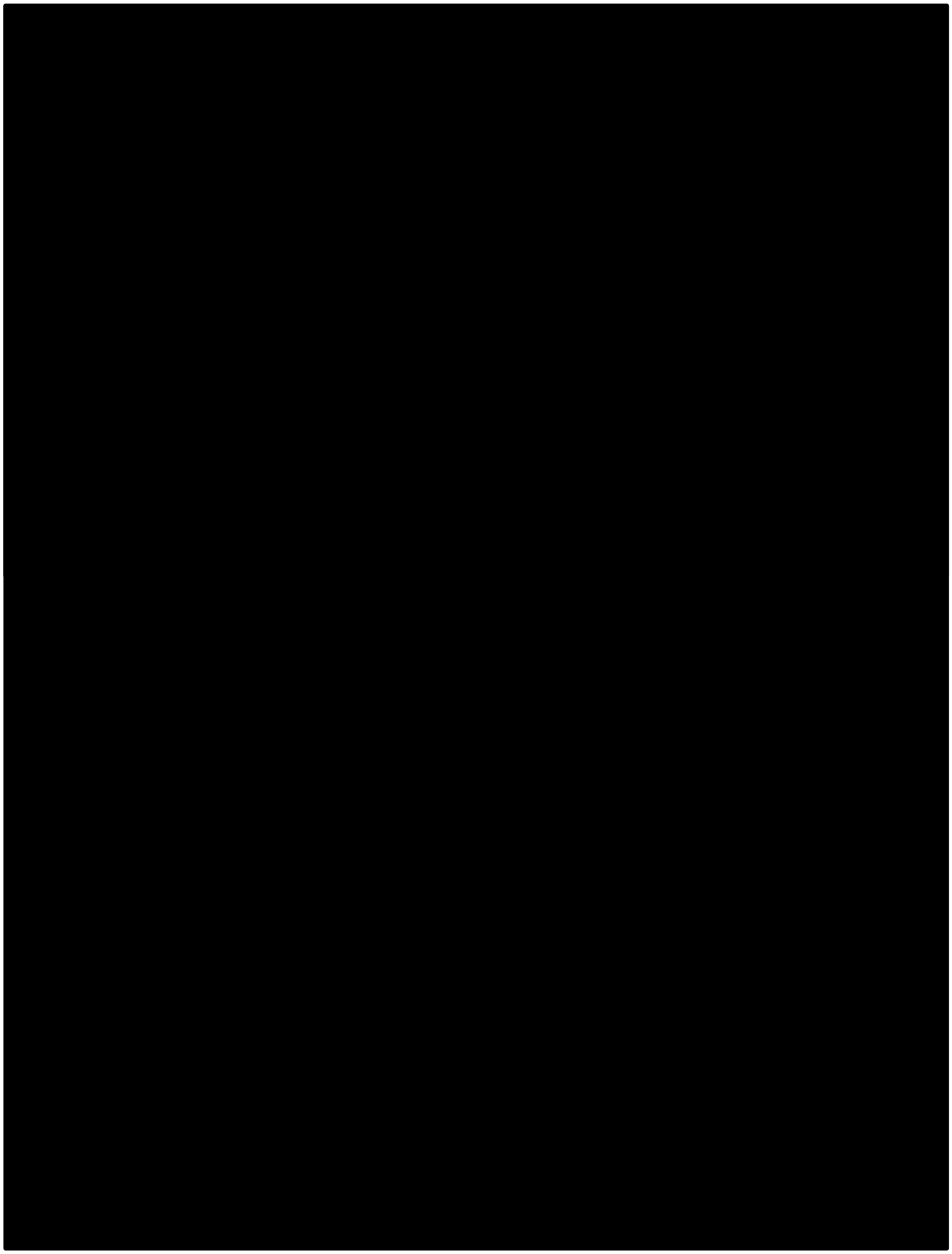
B.2.7.2 Overall Survival

At the time of both CCO1 and CCO2, the majority of patients were alive. Therefore, there were not enough events to make meaningful conclusions for the OS subgroups. Forest plots of the supportive subgroup analyses of OS are found in

Figure 16 (CCO1) and Figure 17 (CCO2).

Unlike the PFS the hazard ratios for CCO1 compared to COO2 have changed across the whole population and the sub-populations which may be the result of placebo patients crossing over cabozantinib. At COO1 the HR for OS was 0.54 (95% CI: 0.27, 1.11) for the overall population and at CCO2 the HR was (HR = 0.76, 95% CI 0.45, 1.31). This change was generally reflected across all the sub-populations.

Figure 16: CCO1 - Forest plots of subgroup analyses for OS (Unstratified Hazard Ratios, ITT population)

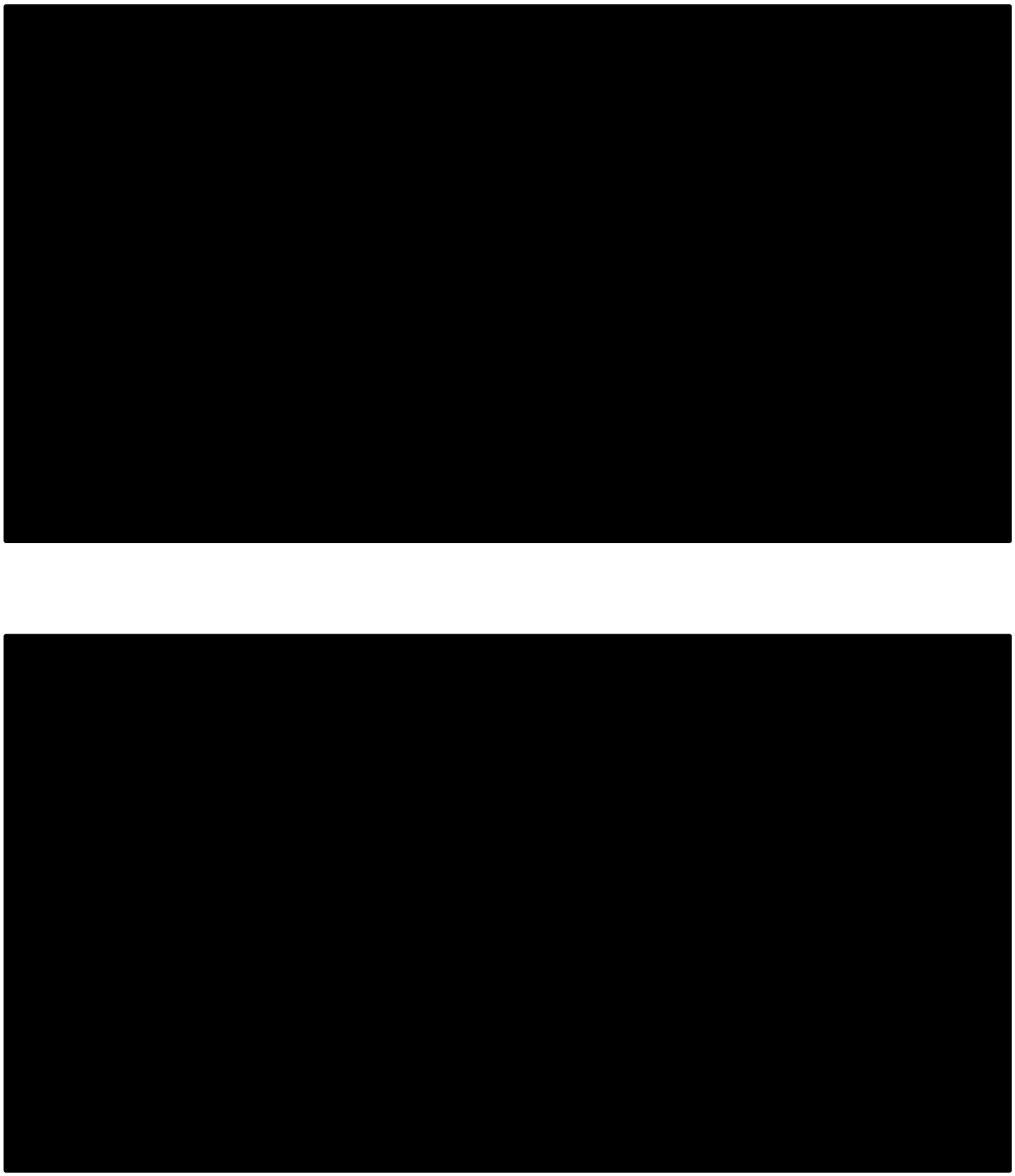


Abbreviations: BIRC – Blinded independent radiology committee; Cabo – Cabozantinib; CI – Confidence interval; CRF – Case report form; DTC – Differentiated thyroid cancer; ECOG PS – Eastern Cooperative Oncology Group

performance status; HR – Hazard ratio; ITT – Intent–to–treat; IxRS – Interactive voice/web response system; NA – Not applicable; NE – Not estimable; PFS – Progression–free survival; RAI – Radioactive iodine; SAP – Statistical analysis plan; VEGFR–TKI – Vascular endothelial growth factor receptor–tyrosine kinase inhibitor; W1D1, Week 1 Day 1.

**Stratification factors are receipt of prior lenvatinib (yes, no) and age at informed consent (≤ 65 years vs > 65 years).
[1] Receipt of prior sorafenib and lenvatinib per CRF
[2] Prior VEGFR-TKI anticancer therapy agents for DTC per subject per history of non-radiation anticancer therapy
[3] ECOG PS at baseline. One subject (3903-3322) in the cabozantinib arm had a predose W1D1 baseline ECOG PS of 2 which was considered the baseline per the SAP. However, the subject had an ECOG PS of 1 at screening.
Source: XL184-311 CSR (30th April 2020)[46]

Figure 17: CCO2 - Forest plots of subgroup analyses for OS (Unstratified Hazard Ratios, ITT population)



Abbreviations: BIRC – Blinded independent radiology committee; Cabo – Cabozantinib; CI – Confidence interval; CRF – Case report form; DTC – Differentiated thyroid cancer; ECOG PS – Eastern Cooperative Oncology Group performance status; HR – Hazard ratio; ITT – Intent–to–treat; IxRS – Interactive voice/web response system; NA – Not applicable; NE – Not estimable; PFS – Progression–free survival; RAI – Radioactive iodine; SAP – Statistical analysis plan; VEGFR–TKI – Vascular endothelial growth factor receptor–tyrosine kinase inhibitor; W1D1, Week 1 Day 1.
**Stratification factors are receipt of prior lenvatinib (yes, no) and age at informed consent (≤ 65 years vs > 65 years).
[1] Receipt of prior sorafenib and lenvatinib per CRF
[2] Prior VEGFR-TKI anticancer therapy agents for DTC per subject per history of non-radiation anticancer therapy
[3] ECOG PS at baseline. One subject (3903-3322) in the cabozantinib arm had a predose W1D1 baseline ECOG PS of 2 which was considered the baseline per the SAP. However, the subject had an ECOG PS of 1 at screening. Source: XL184-311 CSR Addendum 1 (21st May 2021)[51]

B.2.8. Meta-analysis

Meta-analysis is a method of evidence synthesis that combines multiple different independent studies and uses statistical methods to provide an estimate of absolute effect.[59] A phase II trial (ID- NCT02041260) investigated the effects of cabozantinib on RAI-refractory advanced DTC in a first line setting.[60] This trial was conducted in a first line setting so therefore it would not be appropriate to perform a meta-analysis with COSMIC-311. Another phase II trial was a single-arm study (NCT01811212) to assess the efficacy and safety of cabozantinib tablets (60 mg) in 25 adult patients with RAI-refractory DTC after up to two lines of prior VEGFR-targeted therapy.[61] The cabozantinib starting dose was 60 mg/day orally but could be escalated to 80 mg if the patient did not experience a response which is not the licensed regimen for

cabozantinib in DTC.[61] Therefore this trial was not appropriate to perform a meta-analysis with COSMIC-311.

Therefore, COSMIC-311 is the only known trial that investigates the effect of cabozantinib compared with placebo in second line DTC and a meta-analysis is not possible in this context.

B.2.9. Indirect and mixed treatment comparisons

An indirect treatment comparison (ITC) is a method to compare treatments in a similar indication that have a common treatment arm. Lenvatinib and sorafenib have been approved by the European Medicines Agency (EMA) for progressive, metastatic, RAI-refractory DTC. Both drugs target VEGFR and have been investigated in two, large, randomised phase III trials (sorafenib in DECISION and lenvatinib in SELECT).[62,63] Selpercatinib is also indicated at second line for DTC, but it only has a license for a RET positive mutation population prior to receiving one other line of systemic therapy.

In other countries lenvatinib and sorafenib are used as second-line options for RAI refractory DTC, with selpercatinib being approved for this population with a RET mutation. Therefore, on a global level a feasibility assessment (FA) of indirect treatment comparison (ITC) between cabozantinib and other approved treatments of interest (i.e., lenvatinib, sorafenib, and selpercatinib) for RAI-refractory DTC was conducted to determine potential approaches and related limitations. The included studies within the FA assessment were two RCTs that investigated lenvatinib versus placebo (SELECT) and cabozantinib versus placebo (COSMIC-311), and one clinical (phase II) trial, a single arm study that investigated selpercatinib (LIBRETTO-001).

The selected studies were assessed for the feasibility of conducting an anchored matching-adjusted indirect comparison (MAIC). However, the similarity assessments show that no quantitative comparison appears to be feasible through a MAIC and a qualitative comparison would be more appropriate. The qualitative comparison of the SELECT, LIBRETTO-001 and COSMIC-311 trials provides limited information, due to the lack of baseline characteristics and clinical outcomes, specifically for the second-line RAI-refractory DTC patients of the comparator trials.

Furthermore, lenvatinib and sorafenib are not recommended by NICE/NHSE for the second-line treatment of patients and with selpercatinib reimbursed specifically for DTC patients with the RET mutation. Therefore, the FA indicates that best supportive care is the most relevant comparator for cabozantinib in second-line RAI-refractory DTC, in addition to the lack of reimbursed medicines for this indication.

B.2.10. Adverse reactions

B.2.10.1 CCO1

All analyses described here were conducted using the safety population compromised of patients from CCO1 (19th August 2020). The safety profile of cabozantinib in COSMIC-311 was consistent with its known safety profile found in prior studies with single-agent cabozantinib, with no new safety concerns emerging from the study in a RAI-refractory DTC patient population.[64,65] The safety of cabozantinib was assessed in all randomised patients who received any amount of study treatment (either cabozantinib or matched placebo). Analyses based on the safety population were performed according to the actual treatment received. A total of 187 patients (125 in the cabozantinib arm and 62 in the placebo arm) were included.[48] As of CCO1, six (5%) of 125 patients in the cabozantinib group and no patients in the placebo group discontinued treatment due to treatment-emergent adverse events.[48]

Safety assessments included evaluations of adverse events (AEs), serious AEs (SAEs), deaths, clinical laboratory test results, physical examination findings, and vital sign measurements reduced as necessary according to individual tolerability of the study treatment. The median daily dose was 42.0 mg (IQR 32.2-54.5) with cabozantinib and 60.0 mg (52.9-60.0) with placebo;[48] the corresponding median dose intensities were [redacted] and [redacted], respectively. The mean daily dose of all cabozantinib arm was [redacted] mg and [redacted] mg for placebo; corresponding mean dose intensities were [redacted] and [redacted] respectively.[46]

The overall incidence of AEs was 94% in the cabozantinib arm and 93% in the placebo arm. The incidence of treatment-related AEs was higher in the cabozantinib arm (90% versus 52%), and this was also the case with grade three or four AEs (57% versus 26%). There was a low rate of treatment discontinuation from blinded study treatment

due to AEs related to study treatment in each treatment arm (cabozantinib 5%, placebo 0%).[48]

B.2.10.2 CCO2

All analyses in this section were conducted using the safety population comprised of patients from CCO2 (8th February 2021).[66] The incidence and severity of AEs were similar to those found in CCO1, with no important changes observed in cabozantinib safety profile since CCO1. Table 17 gives an overview of the AEs from CCO1 and COO2.

Through CCO2 (8th February 2021), a total of 258 patients (170 cabozantinib, 88 placebo) were randomised to receive study treatment. All randomised patients received study treatment; therefore, the safety and ITT populations are the same. As of CCO2, a total of 137 patients discontinued blinded study treatment, 76 (45%) in the cabozantinib arm and 61 (69%) in the placebo arm. The median daily dose of all cabozantinib arm was 39.5 mg cabozantinib and 60 mg for placebo; the corresponding median dose intensities were [redacted] and [redacted], respectively. The mean daily dose of all cabozantinib arm was [redacted] mg and [redacted] mg for placebo; corresponding mean dose intensities were [redacted] and [redacted], respectively.

The overall incidence of AEs was 98% in the cabozantinib arm and 85% in the placebo arm. The incidence of treatment-related AEs was higher in the cabozantinib arm (94% versus 47%), and this was also the case with grade three or four AEs (62% versus 28%).[66] AEs leading to dose modification (reduction or interruption) were observed in [redacted] and [redacted] of patients in the cabozantinib arm and the placebo arm, respectively. The discontinuation rate of cabozantinib or placebo due to treatment related AEs was 8.8% (15 patients) in the cabozantinib arm and 0% in the placebo arm.[66]

The most common grade 3 or 4 AEs in the cabozantinib group were hypertension (12% vs. 2.3% with placebo), palmar-plantar erythrodysesthesia (10%, vs. 0% with placebo), fatigue (8.8% vs. 0% with placebo), and diarrhoea (7.6% vs. 0% with placebo), and hypocalcaemia (7.6% vs. 2.3% with placebo). Table 17 presents an overview of AEs that were reported in COSMIC-311.[66]

Table 17: Overview of AEs (safety population CCO1 and CCO2)

	CCO1		CCO2	
Parameters	**Cabozantinib (N=125) n (%)**	**Placebo (N=62) n (%)**	**Cabozantinib (N=170) n (%)**	**Placebo (N= 88) n (%)**
Any AE	117 (94)	52 (84)	166 (98)	75 (85)
Treatment-related AE	112 (90)	32 (52)	159 (94)	41 (47)
Grade 3 or 4 AE	71 (57)	16 (26)	106 (62)	25 (28)
Treatment-related Grade 3 or 4 AE	59 (47)	4 (6.5)	[redacted]	[redacted]
Grade 4 AE	7 (5.6)	2 (3.2)	11 (6.5)	2 (2.3)
Treatment-related Grade 4 AE	5 (4.0)	0	[redacted]	[redacted]
Grade 5 AE ≤ 30 days after last dose	[redacted]	[redacted]	[redacted]	[redacted]
Treatment-related Grade 5 AE ≤ 30 days after last dose	0	0	0	0
Treatment-related Grade 5 AE at any time	0	0	0	0
SAE	[redacted]	[redacted]	[redacted]	[redacted]
Treatment-related SAE	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose modification (reduction or interruption)	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose reduction	71 (57)	3 (4.8)	114 (67)	4 (4.5)
AE leading to dose interruption	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to treatment discontinuation (not related to disease under study)	6 (4.8)	0	15 (8.8)	0
Related to study treatment	[redacted]	[redacted]	[redacted]	[redacted]

*Patients are counted only once in each category but may be counted in multiple categories
Abbreviations: AE – Adverse event; SAE – Serious adverse event
For each treatment arm, the frequency and percentage of patients with AEs were tabulated by worst CTCAE grade for overall incidence by system organ class and preferred term or only by preferred term.
Source: XL184-311 CSR Addendum 2 (19th August 2021)[66]

B.2.10.3 Summary of adverse events

Adverse event rates from both CCO1 and CCO2 show that cabozantinib in this indication has a manageable safety profile. The most frequently reported AEs (≥ 20% incidence) in the cabozantinib arm of COSMIC-311 were consistent with the known safety profile of cabozantinib and included diarrhoea, PPE, hypertension, fatigue, ALT increased, nausea, AST increased, decreased appetite, hypocalcaemia, and weight decrease. Grade 3/4 adverse events had a low incidence at approximately 5%. A

summary of all adverse events with a frequency ≥ 10% reported in COSMIC-311 is outlined in Table 18.

Table 18: Summary of Frequent Adverse Events (≥ 10% in Either Treatment Arm; Safety Population, CCO1 and CCO2)

Preferred term	CCO1						CCO2					
	Cabozantinib (N=125) n (%)			Placebo (N=62) n (%)			Cabozantinib (N=170) n (%)			Placebo (N=88) n (%)		
	Any Grade	Grade 3/4	Grade 5	Any Grade	Grade 3/4	Grade 5	Any Grade	Grade 3/4	Grade 5	Any Grade	Grade 3/4	Grade 5
Number of patients with at least one AE	117 (94)	71 (57)	9 (7.2)	52 (84)	16 (26)	1 (1.6)	[redacted]	106 (62)	14 (8.2)	[redacted]	25 (28)	1 (1.1)
Diarrhoea	64 (51)	9 (7.2)	0	2 (3.2)	0	0	[redacted]	13 (7.6)	0	[redacted]	0	0
PPE	57 (46)	13 (10)	0	0	0	0	[redacted]	17 (10)	0	[redacted]	0	0
Hypertension	35 (28)	11 (8.8)	0	3 (4.8)	2 (3.2)	0	[redacted]	20 (12)	0	[redacted]	2 (2.3)	0
Fatigue	34 (27)	10 (8.0)	0	5 (8.1)	0	0	[redacted]	15 (8.8)	0	[redacted]	0	0
ALT increased	30 (24)	1 (0.8)	0	1 (1.6)	0	0	[redacted]	1 (0.6)	0	[redacted]	1 (1.1)	0
Nausea	20 (24)	4 (3.2)	0	1 (1.6)	0	0	[redacted]	4 (2.4)	0	[redacted]	0	0
AST increased	29 (23)	0	0	1 (1.6)	0	0	[redacted]	0	0	[redacted]	0	0
Decreased appetite	29 (23)	4 (3.2)	0	10 (16)	0	0	[redacted]	5 (2.9)	0	[redacted]	0	0
Hypocalcemia	29 (23)	9 (7.2)	0	1 (1.6)	1 (1.6)	0	[redacted]	13 (7.6)	0	[redacted]	2 (2.3)	0
Weight decreased	23 (18)	1 (0.8)	0	3 (4.8)	0	0	[redacted]	4 (2.4)	0	[redacted]	0	0
Asthenia	19 (15)	3 (2.4)	0	9 (15)	0	0	[redacted]	4 (2.4)	0	[redacted]	0	0
Dyspnoea	19 (15)	4 (3.2)	0	11 (18)	2 (3.2)	0	[redacted]	3 (1.8)	0	[redacted]	3 (3.4)	0
Proteinuria	19 (15)	1 (0.8)	0	2 (3.2)	0	0	[redacted]	4 (2.4)	0	[redacted]	0	0
Vomiting	18 (14)	1 (0.8)	0	5 (8.1)	0	0	[redacted]	3 (1.8)	0	[redacted]	0	0
MI	17 (14)	3 (2.4)	0	0	0	0	[redacted]	3 (1.8)	0	[redacted]	0	0
Stomatitis	16 (13)	3 (2.4)	0	2 (3.2)	0	0	[redacted]	6 (3.5)	0	[redacted]	0	0

Hypomagnesaemia	15 (12)	1 (0.8)	0	3 (4.8)	0	0	[redacted]	2 (1.2)	0	[redacted]	0	0
Constipation	13 (10)	0	0	5 (8.1)	0	0	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dysphonia	13 (10)	0	0	1 (1.6)	0	0	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	7 (5.6)	2 (1.6)	0	8 (13)	0	0	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cough	6 (4.8)	0	0	12 (19)	0	0	[redacted]	0	0	[redacted]	0	0
Constipation	NR	NR	NR	NR	NR	NR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dysgeusia	NR	NR	NR	NR	NR	NR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Arthralgia	NR	NR	NR	NR	NR	NR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Headache	NR	NR	NR	NR	NR	NR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypokalaemia	NR	NR	NR	NR	NR	NR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations**:** AE – Adverse event; ALT – Alanine aminotransferase; AST – Aspartate aminotransferase; MI – Mucosal inflammation; PPE – palmar–plantar erythrodysaesthesia syndrome.
At each level of subject summarisation, a subject was counted once for the most severe event if the subject reported one or more events.
Source: XL184-311 CSR (30th April 2020)[46] and Brose et al, 2021[48] and XL184-311 CSR Addendum 2 (19th August 2021)[66]

B.2.11. Ongoing studies

No relevant studies are underway that are anticipated to provide additional evidence within the next 12 months for cabozantinib for the treatment of advanced DTC.

B.2.12. Interpretation of clinical effectiveness and safety evidence

Cabozantinib is indicated as monotherapy for the treatment of RAI-refractory DTC in adults who have progressed on a previous VEGFR therapy. As per ESMO guidelines, sorafenib and lenvatinib are currently the only treatments indicated for untreated RAI-refractory DTC. There is a significant unmet need in post VEGFR treatment for advanced RAI-refractory DTC.[3] The proposed positioning of cabozantinib as a treatment option after progression from sorafenib or lenvatinib treatment will fulfil this unmet need.

Cabozantinib is an oral multi-targeted inhibitor of RTKs that delivers extended survival and significantly delayed disease progression in patients with advanced RAI-refractory DTC who have received prior therapy. This is supported by a robust, high-quality phase III clinical trial.

The COSMIC-311 trial was an international, randomised, double-blinded, 2:1, placebo-controlled, phase III trial. At the cut-off date for the final analysis of the Full ITT population (CCO2 – 8th February 2021) there was high maturity of PFS, with a total of 131 events (either radiographic progression and/or death) with an information fraction of 67.9%. The median time of follow-up through the data cut-off date was 10.1 months.

At CCO2 a lower proportion of subjects in the cabozantinib arm experienced disease progression compared with that in the placebo arm (29% vs 74%, respectively). The analysis demonstrated a statistically significant improvement in PFS for subjects in the cabozantinib arm compared with the placebo arm: the HR, adjusted for stratification factors (per IxRS), was 0.22 (96% CI: 0.15, 0.32; p<0.0001). The KM estimates for median duration of PFS were 11.0 (96%CI: 7.4, 13.8) months in the cabozantinib arm vs 1.9 months (96%CI: 1.9, 3.7) in the placebo arm. The landmark estimate of the

proportion of subjects event-free at 12 months was [redacted] in the cabozantinib arm compared with [redacted] in the placebo arm.

In the Full ITT population (CCO2) the ORR was significantly higher (p=0.0003) for subjects in the cabozantinib arm 11% (95% CI: 6.9, 16.9) versus subjects in the placebo arm 0% (95% CI: 0.0, 4.1). The majority (18/19 subjects) of objective responses in the cabozantinib arm were PRs. There was a higher rate of stable disease (SD) in the cabozantinib arm relative to the placebo arm (68.8% vs 38.6%, respectively). The disease stabilisation rate - DSR (ORR + SD ≥ 16 weeks) was 52.9% (95% CI: [redacted]) in the cabozantinib arm compared with 19.3% (95% CI: [redacted]) in the placebo arm. The frequency of PD as best response was lower in the cabozantinib arm compared with the placebo arm (6.5% vs 47.7%, respectively), indicating a low incidence of primary refractory disease to cabozantinib treatment in this study population. The KM estimate of median (range) duration of objective response (DOR) per BIRC was 10.2 (9.3 to NE) months in the cabozantinib arm. The median (range) time from randomisation to the first objective response per BIRC was 3.6 (1.74, 7.52) months in the cabozantinib arm.

In the Full ITT population (CCO2), a total of 58 deaths (37 cabozantinib, 21 placebo) were reported at CCO2. Survival status as of CCO2 was determined for all 258 randomised subjects. Of note, 133 subjects (78%) in the cabozantinib arm and 67 subjects (76%) in the placebo arm were censored at their last known alive dates including 2 cabozantinib subjects who died after data cut-off.

The analysis of the Full ITT population (CCO2) demonstrated a trend for longer OS for subjects in the cabozantinib arm compared with the placebo arm: the HR, adjusted for stratification factors (per IxRS), was 0.76 (95% CI: 0.45, 1.31). The stratified HR at CCO1 was 0.54 (95% CI: 0.27, 1.11).

For the Full ITT population (CCO2) the KM estimate for median duration of OS was 19.4 months (95% CI: 15.9, NE) in the cabozantinib arm and NE in the placebo arm. Of note, the tail of the KM curve and median estimate for OS are unstable due to the low number of subjects at risk with the longest follow up times. Importantly, the placebo arm included 40 subjects who crossed over to receive cabozantinib, 8 of whom had

an event. The other 32 subjects were censored; of these subjects, 12 had at least 6 months of post-crossover survival and 2 were still on open-label cabozantinib as of CCO2. The placebo crossover subjects were not censored at the time of crossover and were analysed under the randomized placebo arm for OS analysis under ITT principles. At CCO1 the placebo arm included 19 subjects who subsequently crossed over to receive cabozantinib; these subjects were not censored at the time of crossover and were analysed under the randomised placebo arm for OS analysis under intent-to-treat principles.

OS was not a primary endpoint in COSMIC-311. While the trial was not designed to support statistically significant OS, the analysis of both CCO1 and CCO2 supported the trend of longer OS for subjects in the cabozantinib arm compared with the placebo arm, despite crossover between the arms. For CCO2 and CCO1, the stratified HRs were 0.76 (95% CI: 0.45, 1.31) and 0.54 (95% CI: 0.27, 1.11), respectively.

The rate of crossover between the placebo arm and the cabozantinib arm is a significant issue that makes OS results difficult to interpret.

The method of RPSFT was used to adjust OS results for crossover in the trial. All statistical methods, including the RPSFT, that adjust OS for crossover in trials come with assumptions that if not met make the output of these methods subject to bias. The “common treatment affect” assumption in the RPSFT method means that the time when control arm patients cross does not have an effect on the treatment effect and that patients experience the same treatment affect as those who were originally in the treatment arm. This is may not hold as progression of disease in the placebo arm occurred early on in the trial and therefore those patients that crossed over may have had a worse prognosis than those who only received cabozantinib and therefore the OS benefit of cabozantinib is underestimated despite efforts to adjust for crossover. It could be argued that the HR for OS from COO1 could be a better reflection of the survival benefit of cabozantinib as less patients had crossed over.

HRQoL was collected in the trial using EQ-5D-5L. Unfortunately the HRQoL results are only reported for CCO1 and EQ-5D-5L questionnaires were discontinued post progression and for patients who transitioned to the crossover phase. For the data that is available on all dimensions of the EQ-5D-5L, changes from baseline in patients in the cabozantinib and in the placebo arms did not show any statistically or clinically meaningful treatment difference indicating treatment with cabozantinib did not result in a deterioration in QoL to AE's. After week 33, EQ-5D-5L questionnaires were collected in less than 5 patients in the placebo arm. Therefore, it is difficult to interpret results post week 33. Overall, the treatment of RAI refractory with cabozantinib has not shown a quality-of-life deterioration compared to placebo.

The benefits of cabozantinib were accompanied by a manageable safety profile. Patients in the cabozantinib arm (cabozantinib only and those who crossed over) showed a significantly longer duration of exposure to treatment. The median duration of exposure (including dose interruptions) was longer in the 'all cabozantinib' arm compared with the placebo arm (5.5 months vs 2.6 months, respectively) and 6.0 months in the cabozantinib only arm. The rate of treatment discontinuation due to treatment-related AEs in the cabozantinib arm was reasonably low at 8.8% (15 subjects). The most frequently reported AEs in the cabozantinib group were typical of those with VEGFR-TKI therapies, such as sorafenib and lenvatinib and consistent with the known safety profile of patients treated with cabozantinib in other disease areas.[63,67,68]

In conclusion, cabozantinib demonstrated a clinically meaningful and statistically significant prolongation of PFS in patients with progressive RAI-refractory DTC who had previously received a VEGFR-targeted therapy. The benefit was maintained across all prespecified subgroups, including those defined by age, prior receipt of lenvatinib, sorafenib, or both agents. The disease stabilisation rate and consistent reduction in target lesion size all favour treatment with cabozantinib and support the PFS results. Cabozantinib was well tolerated with a manageable and known side effect profile and this was achieved without a detriment to QoL. Analysis of OS in the ITT population demonstrated a clear trend for improvement with cabozantinib but the true

benefit has been confounded by crossover of patients from placebo to active treatment in the trial.

B.3. Cost effectiveness

B.3.1. Published cost-effectiveness studies

An SLR was undertaken on the 14th October 2021 to identify published cost-effectiveness studies relevant to the decision problem (see Section B.1.1).

Please see Appendix G for the methods used to identify all relevant studies, in addition to a description and quality assessment of the cost-effectiveness studies identified.

In line with guidance from the Centre for Reviews and Dissemination (CRD)[69], the population, interventions, comparators, outcomes and study type (PICOS) principal was used to define the following review question to identify relevant cost-effectiveness studies:

- What is the cost-effectiveness of treatments available for RAI-refractory DTC?
    - Which types of economic models have been developed for RAI-refractory DTC?
    - What is the design of these models?
    - What model assumptions were made?
    - What input data (e.g. costs, utilities) was used for these models?
    - What are the cost/quality-adjusted life year (QALY) and cost/life year gained (LYG) results?
    - What are the quality and limitations of the included studies based on NICE recommended quality assessment checklist (i.e. Drummond Checklist)?

Overall, six relevant cost-effectiveness publications were identified based on the selection criteria (See Table 19). All cost utility analyses (CUAs) used a Markov model approach, while the cost effectiveness analyses (CEAs) used partitioned survival models (PSMs) to evaluate the cost-effectiveness of interventions in RAI-refractory DTC (RR-DTC). Five of the models assessed the cost-effectiveness of sorafenib and/or lenvatinib, and the remaining model compared larotrectinib with sorafenib.

Carlson et al.[70] published a PSM to compare the expected life years (LYs) and QALYs for NTRK-positive thyroid cancer patients eligible to receive larotrectinib, sorafenib, or lenvatinib. Although the model did not provide any cost inputs, it was considered relevant for inclusion given the availability of LYs and QALYs gained for the interventions of interest.

The model published by Huang et al.[71] evaluated the cost-effectiveness of lenvatinib and sorafenib. For both analyses, placebo was used as comparator. As the results are only published in a conference abstract, limited information on model structure, perspective and country is available.

Wilson et al.[72] published a Markov model to evaluate the cost-effectiveness of lenvatinib versus sorafenib and placebo, and sorafenib versus placebo in the RAI-refractory DTC population. The model used three health states: stable disease, progressed disease and death with cycle lengths of two months. For analysis, a US "limited" societal perspective was considered to estimate the effect over a life-time horizon.

Carrasquilla-Sotomayor et al.[73] also published a Markov model for the evaluation of sorafenib versus BSC in RAI-refractory DTC. The analysis was conducted from a Colombian perspective and ICERs were expressed in Colombian Peso. No further details were available regarding model structure.

Trembley et al.[74] also conducted a direct comparison between active ingredients (i.e., lenvatinib versus sorafenib). This analysis was performed from a US perspective. No details were available regarding model structure. A 10-year time horizon was chosen to evaluate the cost-effectiveness of lenvatinib, both in terms of cost/QALY as for cost/LYG.

In the model published by Erdal et al.[75] a PSM was used to evaluate the cost-effectiveness of sorafenib versus BSC. The model used three health states: progression-free, progression and death. Costs and effects were evaluated over a 30-years' time horizon using cycle lengths of 28 days. The analysis was conducted from

a Turkish payer perspective, although the ICER was expressed in USD dollar, all costs were calculated in Turkish Liras (TL) and converted to USD.

Table 19: Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Carlson, 2021[70]	2021	Cost effectiveness analysis, Partitioned survival model	NTRK-positive thyroid cancer patients eligible to receive larotrectinib, sorafenib, or lenvatinib.	Larotrectinib: 4.03 Sorafenib: 3.15 Larotrectinib incremental over sorafenib: 0.88	NR	NR
Huang, 2016[71]	2016	Cost utility analysis, US, pairwise comparison	RAI-refractory DTC	No quality-of-life data available	Cost in US dollars	Total cost per QALY; Lenvatinib compared with placebo $95,695
Wilson, 2017[72]	2017	Cost utility analysis, Markov model, Three health states including stable disease, progressed disease and death, two-month cycle length, US, Limited societal perspective, Lifetime horizon	RAI-refractory DTC	Lenvatinib:1.34 Sorafenib: 0.96 Levantinib incremental over sorafenib: 0.38	Total cost in US dollars Lenvatinib: $165,487 Sorafenib: $155,948	Total cost per QALY; Lenvatinib compared with sorafenib $25,275
Carrasquilla-Sotomayor 2017[73]	2017	Cost utility analysis, Markov model, Colombia.	RAI-refractory DTC	No quality-of-life data available for sorafenib or BSC Sorafenib incremental over BSC: 0.67	Cost in Columbian Peso (COP$) per month Sorafenib: $9,138,752 BSC $20,510,821	Total cost per QALY; Sorafenib compared with BSC $16,973,237
Tremblay, 2016[74]	2016	Cost effectiveness analysis, 10-year time horizon, US perspective	DTC	No quality-of-life data available for	Cost per day in US dollars	Total cost per QALY; lenvatinib compared with

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
				larotrectinib or sorafenib Larotrectinib incremental over sorafenib: 0.55	Larotrectinib $438 Sorafenib: $411 Larotrectinib incremental cost over sorafenib: $27	sorafenib $103,925
Erdal, 2015[75]	2015	Cost-effectiveness analysis, partitioned survival model, Turkish payer perspective, 30-year time horizon. Health states include: PFS, progression, Death.	DTC	No quality-of-life data available Sorafenib incremental over BSC: 0.8	All costs were calculated in Turkish Liras (TL) and converted to USD. Sorafenib incremental cost over BSC: $24,384	Total cost per QALY sorafenib compared with BSC $30,485

Abbreviations: BSC – Best supportive care; DTC – Differential thyroid cancer; ICER – Incremental cost-effectiveness ratio; RAI – Radioactive iodine; NTRK –Neurotrophic tyrosine receptor kinase; PFS – Progression free survival; QALYs – Quality-adjusted life years; USD – United States dollars.

B.3.2. Economic analysis

The aforementioned economic SLR (Section B.3.1), identified six relevant economic models for treatment of DTC in adults. All models, whether Markov or PSMs, included the health states progression free survival, progressed disease and death to evaluate the cost-effectiveness of interventions in RAI-refractory DTC. Furthermore, the TA535[76] appraisal for lenvatinib and sorafenib included these health states in their appraisal for the same patient population in first-line treatment for DTC.

For advanced or metastatic cancers, the PSM approach is the most commonly used modelling approach to capture the progressive nature of the condition and is a well-established model framework to assess the cost-effectiveness of oncology treatments. This is especially true in the case of treatments for advanced or metastatic cancers, primarily because they often easily reproduce the observed survival outcomes (i.e., high face validity). The PSM approach allows utilisation of independent overall survival and progression free survival curves from the COSMIC-311 trial constructed from the Kaplan-Meier data. Similarly, previous appraisals in DTC utilised a PSM approach; TA535[55,76] for first-line treatment, TA742[77] for second-line treatment in advanced thyroid cancer (including DTC) with RET alterations and TA630[78] for treating NTRK fusion-positive solid tumours that could include thyroid cancer, and is further supported by the NICE DSU guidance for use within NICE oncology models.[79] No existing economic evaluations of cabozantinib were identified in the cost-effectiveness SLR (Published cost-effectiveness studies), therefore a de novo CEM was developed.

The following sections describe the de novo CEM in depth, including the patient population, model structure, intervention and comparators included in the analysis.

B.3.2.1 Patient population

The population entering the CEM are adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed during or after prior systemic therapy. The population is in line with the EMA and MHRA approval for cabozantinib[1,14] and is the same as the ITT population of the COSMIC-311 phase-3 clinical trial.[46,47] Key inclusion criteria are:

- Age ≥ 18 years old
- DTC patients who are refractory or ineligible to receive RAI therapy (i.e., RR-DTC patients)
- Previously treated with at least one of the following VEGFR-targeting TKI agents for DTC: lenvatinib or sorafenib

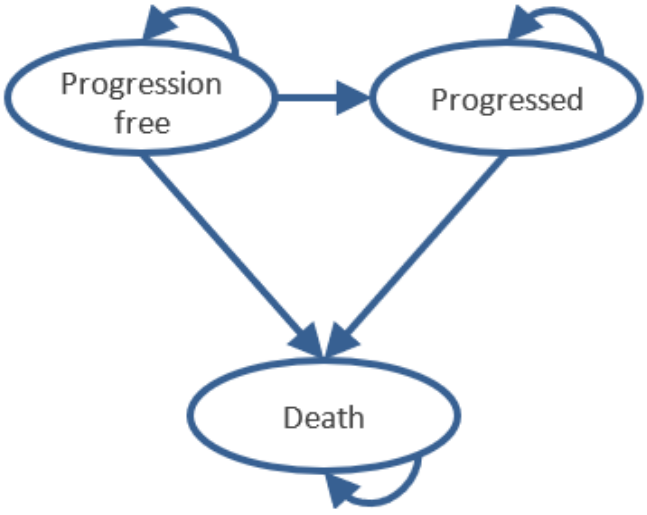
This cohort is reflective of the licenced indication for cabozantinib and the scope for this NICE appraisal and decision problem.

B.3.2.2 Model structure

A PSM over a patient's lifetime was deemed most appropriate to estimate the cost-effectiveness of cabozantinib for the treatment of adults with DTC who have progressed during or after prior systemic therapy. Previous NICE submissions of sorafenib and lenvatinib in RAI-refractory DTC[76] have used the PSM approach, as well as Carlson et al. 2021[70]. In addition, the structure (Figure 18) and health states are consistent with the natural disease progression in oncology patients.[80]

Since the mean age of patients in the COSMIC-311 study was 65 years, a time horizon of 35 years was chosen – assuming no patients survive beyond a mean age of 100 years. The CEM was constructed in Microsoft Excel Office 365. All selected model inputs and rationale are displayed in Table 20 below.

Figure 18: Model schematic for partition survival model



The PSM includes three mutually exclusive health states: progression free, progressed disease and death. All patients enter the model in the progression free state. Patients either stay in the progression free state, or progress to progressed disease state or death and cannot improve their health state. The proportion of patients in each health state at every cycle (month) is estimated directly from the parametric distributions fitted to the PFS and OS data from the COSMIC-311 trial using data from CCO2 (Full ITT population). To account for the cross-over and obtain an unbiased estimate of the OS benefit associated with cabozantinib, the RPSFT adjustment method was used, in line with NICE DSU TSD 16[81], to adjust for this cross-over and estimate the OS associated with the BSC arm (see Section B.2.3.2 and Section B.2.4.2.2).

The time on treatment was determined by the time to treatment discontinuation (TTD) data from the COSMIC-311 trial.

Baseline characteristics are based on data from the COSMIC-311 trial. Cost categories considered in the model include: treatment costs; health state costs and adverse event costs. Utility values for the health states were based on UK clinicians validated[50] values published by Fordham et al. 2015[82] as they commented that a 0.35 decrement from PFS to PD would be expected, and that other sources produced implausible values (Section B.3.4).[46,55]

Costs and health-related utilities were allocated to each health state and multiplied by state occupancy to calculate the weighted costs and QALYs per cycle. Effectiveness measures included life years (LYs) and QALYs. The incremental cost-effectiveness ratio (ICER) of cabozantinib versus BSC was evaluated in terms of the incremental cost per QALY gained.

The analysis was conducted from the perspective of the NHS, including direct medical costs and Personal Social Services (PSS) costs over a lifetime time horizon of the patient cohort from the initiation of treatment. A monthly cycle length was considered in the base case, and both costs and effects were discounted at 3.5% annually. The economic analysis is conducted using the most recent estimates of resource use and

treatment costs available from published sources (2020/21). Costs quoted for other cost-years are inflated to the model cost-year as applicable.

Table 20: Features of the economic analysis

	Current evaluation	
Factor	**Chosen values**	**Justification**
Cycle length	1 month (30.44 days)	The monthly cycle (30.44 days=365.25/12) captures all relevant costs and health outcomes and is consistent with previous technology appraisals for DTC.[55] Shorter cycle lengths may overcomplicate the model calculation given the lifetime horizon of 30 years. Whereas longer cycle lengths increase the risk of over or under predicting costs per QALYs when averaging across cycle times.
Perspective	NHS/PSS	NICE reference case.[83]
Model type	PSM	PSM is a well-established model framework to assess the cost-effectiveness of oncology treatments and has been used in many prior NICE submissions, especially in the case of treatments for advanced or metastatic cancers, primarily because they often easily reproduce the observed survival outcomes (i.e., high face validity). The health states are consistent with the natural disease progression in oncology patients.
Time horizon	Lifetime (35 years)	The time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect all important differences in costs or outcomes between technologies being compared.[83] Therefore, a lifetime horizon was chosen since patients accumulate differential costs and QALYs until death. Since the mean age of patients in the COSMIC-311 study was 65 years, a time horizon of 35 years was chosen – assuming no patients survive beyond a mean age of 100 years.
Discounting	3.5%	NICE reference case.[83] The impact of alternative discount rates has been tested in sensitivity analyses.
Treatment waning effect?	N/A	There is no treatment waning effect applied as patients discontinue cabozantinib treatment when they no longer benefit from therapy or until unacceptable toxicity occurs. Therefore, treatment efficacy is assumed to be reflective of that observed in the COSMIC-311 trial.
Source of utilities	Fordham et al. 2015.[82] Ara et al. 2010.[84]	Quality-of-life data were available from the COSMIC-311 trial, however, as outlined in Section B.3.4.1, the utility values produced were inconsistent with previous oncology treatment. Following consultation with three UK clinicians[50], the Fordham et al. 2015[82] utility values for PFS and PD were deemed suitable as it reflects health state utility

		values for patients in RAI refractory DTC. The Fordham et al. 2015 utilities were also accepted in TA742[77] which represented a second-line treatment setting like cabozantinib. Health states utilities are age-adjusted using the age-decrement equation from Ara et al. 2010[84]. Adverse event disutilities were included for the first cycle of the model as we assume patients experience adverse events in the first month following treatment initiation and are resolved with dose interruption. Adverse event rates of grade 3 and above were sourced from the COSMIC-311 trial and disutility values from published literature.
Source of costs	National Schedule of Reference Costs (2020-21).[85] Georghiou T, Bardsley M. Nuffield Trust.[86] BNF costs.[15] PSSRU report 2021.[87]	Where possible, costs were obtained from UK national resources to reflect the UK NHS/PSS perspective. National schedule of reference costs was used to identify cost of resources used by patients based on their health state. Georghiou et al. 2014[86] provides an end of life cost that is applied as a one-off cost to patients who die. PSSRU pay and prices indices were used to inflate costs to 2020/21[87].

Abbreviations: DTC – Differentiated thyroid carcinoma; NHS – National Health Service; NICE – National Institute for Health and Care Excellence; PD – progressed disease; PFS – Progression free survival; PSM – Partitioned survival model; PSS – Personal Social Services; PSSRU – Personal Social Services Research Unit; QALY – Quality adjusted life year.

B.3.2.3 Intervention technology and comparators

The intervention is cabozantinib and is administered as a 60 mg oral tablet once per day. Cabozantinib is indicated for the treatment of adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI, who have progressed during or after prior systemic therapy, aligned with the population in COSMIC-311.[46] For more information on the product characteristics of cabozantinib, please see Appendix C.

As discussed in Section B.1.2, no treatment is currently recommended by NICE for adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI, who have progressed during or after prior systemic therapy apart from selpercatinib for RET-fusion thyroid cancer (which can include DTC). Selpercatinib is not a comparator in this appraisal as described in Section B.1.1.2. Therefore, only BSC has been included in the model as a comparator in the base case analysis. Patients on BSC do not receive any active treatment regimen and placebo data from COSMIC-311 trial[46] will be used to inform the BSC arm.

B.3.3. Clinical parameters and variables

The primary source of survival data was the COSMIC-311 trial.[46] The proportion of patients and time spent in each health state for the PSM were derived based on the area under the survival curves. Effectiveness inputs for the Full intention-to-treat (ITT) population (Section B.2.4.1), using the latest data cut-off date on the 8th of February 2021 (CCO2), are described in this section.

B.3.3.1 Baseline demographics

Baseline demographics for the modelled cohort were based on the Full ITT population in COSMIC-311 trial[46,47] (see Table 21).

Table 21: COSMIC-311 baseline patient demographics

Baseline demographics	Full ITT	Reference
Mean age (years)	65.0	COSMIC-311[46,47]
% Male	47%	

Abbreviations: ITT – Intention-to-treat

B.3.3.2 Progression free survival

As described in Section B.2.6.1, the COSMIC-311 study met the primary endpoint of PFS at the prespecified interim analysis at CCO1 and then enrolment stopped. There was a second later analysis point, CCO2, which had a median follow-up of 10.1 months. As the follow-up period for PFS was shorter than the model lifetime horizon, extrapolation from the PFS data was required. As recommended in the NICE DSU TSD 14[88], standard parametric models, including the exponential, Weibull, lognormal, log-logistic, Gompertz, and generalized gamma, were fitted to PFS data from COSMIC-311 trial.

The model was selected based on the goodness of fit (AIC and BIC), visual inspection against the observed KM data, and three UK clinicians[50] in an advisory board held in August 2022 inspected whether the extrapolations were clinically and biologically plausible. Figure 19 and

Figure **20** show PFS data fitted and extrapolated using the standard parametric models for the cabozantinib and BSC arms, respectively. The AIC and BIC values for the models are presented in Table 23. These indicate that the log-logistic is the best-fitting model for the cabozantinib arm. For the BSC arm, the best fitting models were the generalized gamma and the log-normal, according to AIC and BIC, respectively.

Based on visual inspection against the observed KM data, proportion of individuals progression-free at landmark timepoints (Table 22) and inspection by three UK clinicians[50] in an advisory board, the Weibull and Gompertz were both deemed clinically plausible. However, coupled with goodness of fit statistics, the Weibull distribution was recommended and has been selected to extrapolate the PFS data in the base case of the model for cabozantinib and BSC. As shown in Table 22, the Weibull estimates [redacted]% in the cabozantinib arm are progression-free at three months dropping to [redacted]% at one year. For the BSC arm, the Weibull estimates [redacted]% at three months dropping to [redacted]% after one year. Upon applying PFS in the model a rule was also applied whereby the PFS curve could not exceed the OS curve for each treatment.

Table 22: Proportion of individuals progression-free in the cabozantinib and BSC arms

Abbreviations: BSC – Best supportive care; Cabo – Cabozantinib; PFS – Progression-free survival
* This distribution was selected as the best fitting model, based on minimisation of AIC, visual assessment, and clinician validation, with which to extrapolate the COSMIC-311 PFS data – CCO2 Full ITT population

Figure 19: PFS curves for Cabozantinib based on COSMIC-311

Figure 20: PFS curves for BSC based on placebo arm of COSMIC-311

Table 23: Parametric Survival models AIC and BICs for PFS based on COSMIC-311

[redacted]	[redacted]			[redacted]			[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

* This distribution was selected as the best fitting model, based on minimisation of AIC, visual assessment, and clinician validation, with which to extrapolate the COSMIC-311 PFS data – CCO2 Full ITT population

Abbreviations: AIC – Akaike's Information Criterion; BIC – Bayesian Information Criterion; BSC – Best supportive care; PFS – Progression free survival

B.3.3.3 Overall survival

The COSMIC-311 trial was designed to allow cross-over at the time of BIRC-confirmed progression (e.g., patients may be switched from BSC to cabozantinib treatment upon disease progression). To account for the cross-over and obtain an unbiased estimate of the OS benefit associated with cabozantinib, the RPSFT adjustment method was used, in line with NICE DSU TSD 16[81], to adjust for this cross-over and estimate the OS associated with the BSC arm (see Section B.2.3.2 and Section B.2.4.2.2).

B.3.3.3.1. RPSFT methodology

The RPSFT uses a counterfactual framework to estimate the survival time gained or lost by receiving active treatment, where counterfactual survival times refer to those that would have been observed if no treatment had been given. In essence, the models assume that active therapy is acting on mortality by multiplying survival by a certain factor (treatment effect) once a patient starts receiving the treatment. This factor may be interpreted as the increase or decrease in survival by taking the active treatment compared to the control treatment. Once established, the survival duration of patients is reconstructed, and re-censored as if they had never received the active compound.[88]

B.3.3.3.2. Analysis

All RPSFT analyses were conducted using the rpsftm-an R package version 1.2.7 for rank preserved structural failure time models.[89] The package allows estimating the treatment effect, θ, using a g-estimation procedure to find the value of θ such that a test statistic $Z(\theta) = 0$. The default test is the log rank test, but alternatively the Wald test from a Cox regression model and a Weibull AFT model can be used by specifying test=coxph or test=survreg. In the current analysis, the coxph test was used to estimate the treatment effect.

The “naive” 95% confidence interval (CI) for the relative treatment effect (i.e. Hazard Ratio) was estimated from the Cox models. However, these intervals could be biased due to artificial censoring in the structural model. Unbiased CIs of the HR were computed by inflating the standard error of the log-hazard ratio to preserve the ITT p-

value.[90] In this analysis, inflated 95% CI of the HR using the latter method were provided.

Furthermore, a sensitivity analysis to test the common treatment effect assumption was conducted. The RPSFT model was rerun for a range of value *k*, where it was assumed that the treatment effect in the placebo is *k* times the treatment effect in the cabozantinib arm.

Prior to treatment cross-over adjustment, an HR of 0.76 (95% CI: 0.45, 1.31) was estimated between cabozantinib and placebo using the original Full ITT analysis (CCO2) (Section B.2.6). After adjusting placebo for treatment crossover using RPSFT method with the common treatment effect assumption, the stratified HR was estimated at 0.65 (0.28, 1.53) for cabozantinib vs placebo (Table 24).

Table 24: HRs (95% CI) for Cabozantinib vs placebo-RPSFT adjusted OS data

Parameter	Cabozantinib	Placebo	Placebo-adjusted
Events, n (%)	37 (22)	21 (24)	21 (24)
Unstratified HR	NA	0.78 (0.45, 1.33)	0.67 (0.39, 1.15)
Stratified HR (naive 95% CI)	NA	0.76 (0.45, 1.31)	0.65 (0.38, 1.13)
Stratified HR (inflated 95% CI)*	NA	0.76 (0.45, 1.31)	0.65 (0.28, 1.53)

*The naïve HR does not consider the dependencies of the data, thereby it underestimates the CI. In this HR, the CI is inflated to account for the underestimation.
Abbreviations: CI – Confidence interval; HR – Hazard ratio; OS – Overall survival; RPFST – Rank preserving structural failure time.

The standard parametric curves for OS for cabozantinib and BSC following adjustment are presented in Figure 21 to

Figure **26**. Four out of the six distributions (Weibull, Gompertz, Log-logistic and Generalised Gamma) unrealistically show that the cabozantinib and BSC curves cross. This has been validated as unrealistic by clinicians at an advisory board held in August 2022.[50]

Table 26 presents the AIC and BIC values for the models. According to the AIC and BIC of the curves which do not cross (Exponential and Lognormal), the Exponential was the best fitting model for cabozantinib and Log-normal for BSC adjusted using RPSFT.

Table 25 displays the survival estimates for cabozantinib and BSC at different timepoints and distributions. UK clinicians[50] advised the survival estimates for the extrapolations for BSC at 5 and 10 years were overestimated and that 0% of patients would be expected to be alive at 5 years although one clinician did think maybe 1% could be alive at 5 years. Of the plausible Exponential and Lognormal curves, the Exponential has a sharper decline in the proportion of patients alive over time and was deemed most appropriate with a cap for BSC applied at five years. Therefore, the Exponential was used in the base case to model the OS of patients receiving cabozantinib and BSC with 0% of BSC patients modelled as being alive from five years.

Table 25: Proportion of individuals alive in the cabozantinib and BSC arms

[redacted]

Abbreviations: BSC – Best supportive care; Cabo – Cabozantinib; NC – Not computed; PFS – Progression-free survival
* This distribution was selected as the best fitting model, based on minimisation of AIC, visual assessment, and clinician validation, with which to extrapolate the COSMIC-311 OS data – CCO2 Full ITT population with RPSFT adjustment.
^Converged above 50%, so the equation cannot calculate 50% median.

Figure 21: Weibull overall survival data

Figure 22: Gompertz overall survival data

Figure 23: Log-logistic overall survival data

Figure 24: Generalised gamma overall survival data

Figure 25: Exponential overall survival data

Figure 26: Lognormal overall survival data

Table 26: Parametric survival models AICs for COSMIC-311 OS

[redacted]	[redacted]			[redacted]			[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

* This distribution was selected as the best fitting model, based on minimisation of AIC, with which to extrapolate the COSMIC-311 OS data – CCO2 Full ITT population with RPSFT adjustment

Abbreviations: AIC – Akaike's Information Criterion; BIC – Bayesian Information Criterion; BSC – Best supportive care; RPFST – Rank Preserving Structural Failure Time.

B.3.3.4 Time to treatment discontinuation

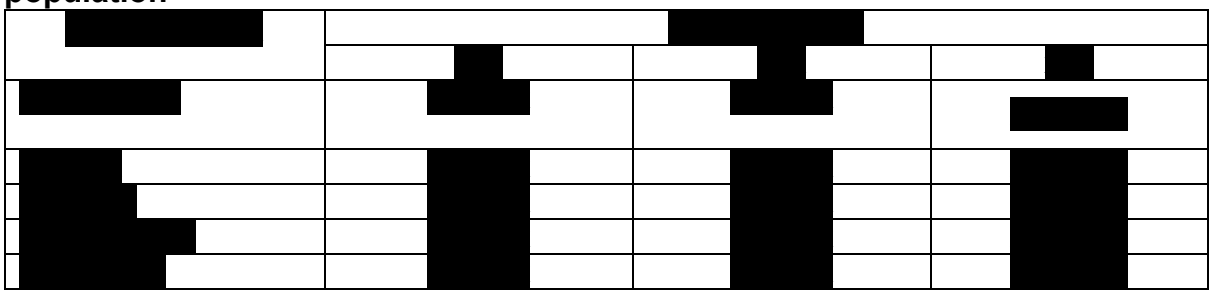
Standard parametric models were also fitted to the TTD data obtained from the COSMIC-311 trial to extrapolate the TTD beyond trial duration.

Figure 27 illustrates the TTD data observed in COSMIC-311 trial and extrapolated using the standard parametric models for the cabozantinib arm. For BSC arm, no TTD data were used as patients in BSC arm are not receiving any active treatment.

Table 27 presents the AIC and BIC values for the models and indicates that the exponential was the best fitting models for the cabozantinib arm and therefore used in the base case. Upon applying TTD in the model a rule was also applied whereby the TTD curve could not exceed the PFS curve for cabozantinib since patients are assumed to discontinue treatment upon progression as per the SmPC.[14]

Figure 27: Cabozantinib TTD data fitted and extrapolated using standard parametric models (CCO2 full ITT population)

Table 27: Parametric survival models AICs for COSMIC-311 TTD data – ITT population



	[redacted]			[redacted]			[redacted]			[redacted]	

*This distribution was selected as the best fitting model to extrapolate the COSMIC-311 TTD data – CC02 Full ITT population
Abbreviations: AIC – Akaike's Information Criterion; BIC – Bayesian Information Criterion; ITT –Intention-to-treat; TTD – Time to treatment discontinuation.

B.3.4. Measurement and valuation of health effects

B.3.4.1 Health-related quality-of-life data from clinical trials

The EQ-5D-5L data collected within the COSMIC-311 trial[46] was analysed to estimate health state utility values. The EQ-5D instrument has the following five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.[91] Each dimension is rated on a five-point scale from 1 (no problem) to 5 (extreme problem). In addition, EQ-5D contains a graded vertical visual analogue scale (1-100) to rate patient's general health state at the time of assessment. In the COSMIC-311 trial, EQ-5D responses were provided by patients at various assessment/time points. For those patients who crossed over treatment, utility assessments were discontinued. As a result, it was not possible to obtain post-crossover specific utility values.

Linear mixed-effect models were used to derive health state utility values ranging from 0 to 1. Several model structures were considered, including random intercepts, random slopes, and random intercepts and slopes. Several potential covariates were included in the models, such as age, gender, treatment arm, assessment time points, progression state. The preferred model structure included a random intercept at the subject level. In addition, only binary indicators for the current progression state were statistically significant in the model. Importantly, the treatment arm was not found to be a statistically significant variable, indicating that one health state-specific utility value across both treatment arms could be used in the cost-effectiveness model.

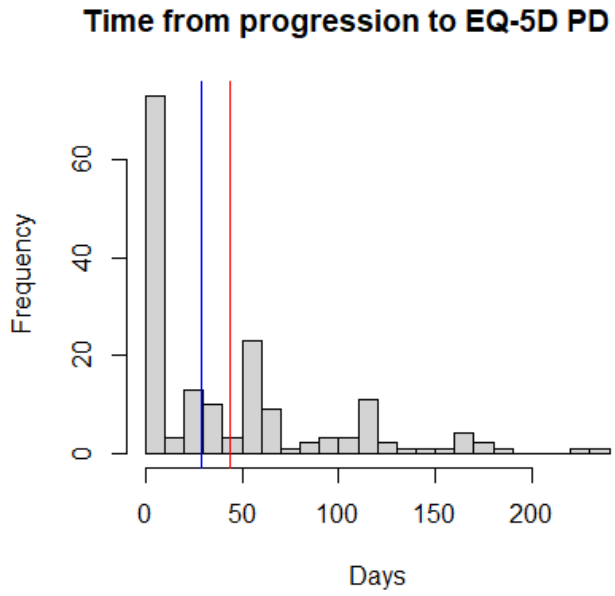
The EQ-5D-5L data from the COSMIC-311 trial was mapped to the EQ-5D-3L using the cross-walk approach by Hernandez-Alava and Pudney (2017)[92] as dictated by recently published NICE guidelines (2022)[83].The health state utility values from the COSMIC-311 analysis are [redacted] for PFS and [redacted] for PD. In addition, the utility value of patients in the death state is assumed to be zero, as per standard convention.

However, the limited impact on utility associated with progression does not appear to be consistent, given the difference between PFS and PD states observed in other models and appraisals in advanced thyroid cancer, this inconsistency was also validated by UK clinicians in a recent advisory board.[16,32–36,50] For example, health state utility values from the DECISION trial of sorafenib in a first-line setting (measured using the EQ-5D-3L) used in multiple-technology appraisal (MTA) (TA535) by the assessment group were 0.72 and 0.80 for patients in PFS receiving sorafenib and BSC, respectively.[76] While individuals in PD state had a utility of 0.64. This equates to a much larger impact associated with progression than that observed in the utility analyses of the COSMIC-311 data. Also, a vignette study by Fordham et al. 2015[82], which aimed to estimate health state utilities in individuals with RR-DTC has also been used and accepted in several NICE appraisals in this clinical area, including TA742[77] in a second-line setting and TA516[93]. In this study, utilities of 0.87 and 0.52 were estimated for the PFS and PD states, respectively.

The limited impact of progression in the COSMIC-311 data was likely a result of limited follow-up in the PD state or missing data, as the data suggests that utility falls over time in the PD state. Regarding missing data, the CSR states that 115 progression events occurred before the data cut-off, however only 89 participants are captured in the HRQoL assessment after progression. This is due to the fact that HRQoL assessment were discontinued in patients who progressed in the placebo arm and began crossover cabozantinib treatment (n=40). Among those captured within the PD HRQoL data, if those in worse health are more likely to drop-out of HRQoL assessments while in the PD state, this could overestimate the progressed disease value. Figure 28 shows the time between progression and HRQoL assessments in the PD state. The median number of days between progression and HRQoL assessment was 29 days, with a mean of 43.7 days. However, the histogram in Figure 28 shows that a large number of PD observations were captured within 10 days of progression (n=73; 43.5% within the first 10 days and n=62; 36.9% within the first 5 days). If the impact of progression on HRQoL is not immediately felt and increases over time, it is unlikely that the PD utility values obtained from this data will be reflective of the full PD state. From the data available, a trend towards utility declining over time during progression can be observed (Table 28). Therefore, if the duration of PD follow up was

shorter than patients would be expected to spend in PD in real life, it is likely that the PD value from the COSMIC-311 trial is not fully reflective of the PD state as a whole. Due to this lack of validity of the COSMIC-311 HRQoL data, Fordham et al. 2015[82] utilities were used in the base case (see Section B.3.4.4).

Figure 28: Histogram showing time from progression to assessment for PD EQ-5D observations (obs)



Median	29
Mean	43.7
95% confidence interval	(35.8, 51.7)
25% quartile	1
75% quartile	62.5
Minimum	0
Maximum	231

Table 28: Descriptive statistics for utility values by health state (COSMIC-311)

Health state	Total obs.	Unique subjects	Mean utility (SD)	Standard error	Median utility	Minimum utility	Maximum utility
PFS (baseline measurement)	253	253	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS (all measurements)	1278	256	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PD (first measurement)	89	89	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

PD (last measurement)	89	89	[redacted]				
PD (last measurement, only in those with multiple assessments)	36	36	[redacted]				
PD (all measurements)	168	89	[redacted]				

Abbreviations: Obs.– Observations; PD – Progressed disease; PFS – Progression free survival; SD – Standard deviation

B.3.4.2 Health-related quality-of-life studies

An economic SLR was undertaken on the 14th October 2021 to identify existing studies investigating HRQoL in management of adults with RAI-refractory DTC.

Please see Appendix G and H for the methods used to identify all relevant studies, and description of the HRQoL studies identified.

The review question and sub-questions evaluated in the HRQoL SLR were:

- What are the impacts of RAI-refractory DTC and its treatment on the HRQoL of patients?
  - Which studies investigate the utilities and HRQoL values associated with RAI-refractory DTC?
  - What is the design of these studies?
  - Which HRQoL and utility values have been reported for RAI-refractory DTC patients?

In total, six studies (reported in seven publications) provided insights on HRQoL/utilities of RAI-refractory DTC patients. Only one study provided data specifically on the ≥2nd line RAI-refractory DTC population. Hudgens et al.[94] delineated HRQoL in patients receiving lenvatinib as first-line and second line therapy. All other studies reported HRQoL outcomes for the overall RAI-refractory DTC population. Overall, interventions varied across studies or were not reported due to the nature of the study design. In three studies – Giani et al.[95], Taylor et al.[96]; and Hudgens et al.[94] – patients received lenvatinib (different doses; 18 mg, 24 mg). Raef et al.[97] investigated patients that received any TKI while patients in Ballal et al.[98] received ^{177}Lu-DOTAGA.(SA.FAPi)$_2$. In the study described by Kerr et al.[99] and

Fordham et al.[82] no interventions were involved. Table 29 presents the results of the HRQoL/utilities studies.

Table 29: HRQoL study results

First author, Year	Instrument	Outcome						
		Treatment arm/group	Variable definition	Baseline Mean (SD)	Timepoint of Assessment	Mean (SD) value at follow up	Mean ∆QoL	P-value
Giani, 2021[95]	EORTC-QLQ-C30	Overall population	Global health status/QoL	67.28 (23.56)	Monthly	NR	0.427 (0.83*)	0.608
			Physical functioning	81.6 (20.28)	Monthly	NR	-0.582 (0.88*)	0.511
			Role functioning	80.86 (26.84)	Monthly	NR	-0.242 (1.23*)	0.844
			Emotional functioning	78.7 (15.39)	Monthly	NR	0.761 (0.80*)	0.321
			Cognitive functioning	87.65 (18.83)	Monthly	NR	0.454 (0.792*)	0.567
			Social functioning	82.72 (21.42)	Monthly	NR	-0.42 (1.01*)	0.677
			Fatigue	29.63 (24.65)	Monthly	NR	0.245 (1.035*)	0.813
			Nausea and vomiting	3.7 (8.44)	Monthly	NR	0.548 (0.702*)	0.436
			Pain	22.22 (26.95)	Monthly	NR	-0.345 (1.038*)	0.74
			Dyspnoea	27.16 (26.21)	Monthly	NR	-1.837 (1.044*)	0.08
			Insomnia	18.52 (25.04)	Monthly	NR	-0.173 (0.85*)	0.838
			Loss of appetite	14.81	Monthly	NR	0.114	0.922

				(21.35)			(1.171*)	
			Constipation	18.52 (26.69)	Monthly	NR	-1.052 (0.962*)	0.276
			Diarrhoea	8.64 (14.89)	Monthly	NR	2.253 (0.862*)	0.01
			Financial difficulties	6.17 (16.11)	Monthly	NR	0.244 (1.081*)	0.822
	VAS	Overall population	C30 Numeric pain rating scale	1.59 (2.09)	Monthly	NR	-0.120 (0.110*)	0.277
Kerr, 2014[99]	TTO & VAS	Overall population	Stable disease	0.86 (95% CI: 0.83-0.89)	NA	NA	NA	NA
			Treatment response	0.8 (95% CI: 0.77-0.84)	NA	NA	NA	NA
			Progressive disease	0.5 (95% CI: 0.45-0.56)	NA	NA	NA	NA
			Stable + grade I-II alopecia***	0.75 (95% CI: 0.71-0.79)	NR	NR	NR	NR
			Stable + grade III fatigue***	0.72 (95% CI: 0.67-0.77)	NR	NR	NR	NR
			Stable + grade III Hand Foot Syndrome (HFS)***	0.52 (95% CI: 0.46-0.58)	NR	NR	NR	NR
			Stable + grade III diarrhoea***	0.42 (95% CI: 0.36-0.48)	NR	NR	NR	NR
Fordham, 2015[82]	TTO & VAS (EQ-5D-3L)	Overall population (Observed Utilities:	Base state – stable/no response	0.8	NR	NR	NR	NR

		Mean observed TTO health state utilities)		(0.19; 95% CI: 0.77-0.84)				
			Response to therapy	0.86 (0.15; 95% CI: 0.83-0.89)	NR	NR	NR	NR
			Progressive disease	0.5 (0.28; 95% CI: 0.45-0.56)	NR	NR	NR	NR
			Diarrhoea	0.42 (0.29; 95% CI: 0.36-0.48)	NR	NR	NR	NR
			Fatigue	0.72 (0.24; 0.67-0.77)	NR	NR	NR	NR
			Hand and foot syndrome	0.52 (0.3; 95% CI: 0.46-0.58)	NR	NR	NR	NR
			Alopecia	0.75 (0.21; 95% CI: 0.71-0.79)	NR	NR	NR	NR
		Overall population (Unadjusted Utilities: Derived from reduced parameter model [health states only])	Base state – stable/no response	0.86 (95% CI: 0.83-0.9)	NR	NR	NR	NR
			Response to therapy	0.04 (95% CI: 0.01-0.07)	NR	NR	NR	NR

			Progressive disease	−0.37 (95% CI: -0.43- -0.31)	NR	NR	NR	NR
			Diarrhoea	−0.48 (95% CI: -0.54- -0.43)	NR	NR	NR	NR
			Fatigue	−0.08 (95% CI: -0.13- -0.04)	NR	NR	NR	NR
			Hand and foot syndrome	−0.35 (95% CI: -0.42- -0.29)	NR	NR	NR	NR
			Alopecia	−0.05 (95% CI: -0.09- -0.01)	NR	NR	NR	NR
		Overall population (Adjusted Utilities: Adjusted for educational qualification level and EQ-5D-3L [usual activities and anxiety/depression] ratings using UK norms)	Base state – stable/no response	0.87 (95% CI: 0.84-0.91)	NR	NR	NR	NR
			Response to therapy	0.04 (95% CI: 0.01-0.07)	NR	NR	NR	NR
			Progressive disease	−0.35 (95% CI: -0.41- -0.29)	NR	NR	NR	NR
			Diarrhoea	−0.47	NR	NR	NR	NR

				(95% CI: -0.52- -0.41)				
			Fatigue	−0.08 (95% CI: -12-0.04)	NR	NR	NR	NR
			Hand and foot syndrome	−0.34 (95% CI: -0.40-0.28)	NR	NR	NR	NR
			Alopecia	−0.05 (95% CI: -0.08- -0.01)	NR	NR	NR	NR
Taylor, 2021[96]	EQ-5D VAS	LEN18	NR	NR	NR	NR	-5.68 (1.619*)	NR
		LEN 24	NR	NR	NR	NR	-5.25 (1.601*)	NR
		LEN 18 vs LEN 24	NR	NR	NR	NR	-0.42 (95% CI: -4.880-4.03)	0.8507
	HUI	LEN18	NR	NR	NR	NR	-0.08 (0.018*)	
		LEN 24	NR	NR	NR	NR	-0.06 (0.017*)	
		LEN 18 vs LEN 24	NR	NR	NR	NR	-0.02 (95% CI: -0.07-0.03)	0.4586
	FACIT/FACT instruments	LEN18	Total score	NR	NR	NR	-4.14 (1.348*)	
		LEN 24	Total score	NR	NR	NR	-4.61 (1.397*)	
		LEN 18 vs LEN 24	Total score	NR	NR	NR	0.47	0.8132

							(95% CI: -3.45-4.39)	
		LEN18	Physical well-being	NR	NR	NR	-3.13 (0.518*)	
		LEN 24	Physical well-being	NR	NR	NR	-3.61 (0.51*)	
		LEN 18 vs LEN 24	Physical well-being	NR	NR	NR	0.48 (95% CI: -0.95-1.92)	0.5058
		LEN18	Social/family well-being	NR	NR	NR	-0.07 (0.525*)	
		LEN 24	Social/family well-being	NR	NR	NR	0.03 (0.518*)	
		LEN 18 vs LEN 24	Social/family well-being	NR	NR	NR	-0.1 (95% CI: -1.54-1.34)	0.8886
		LEN18	Emotional well-being	NR	NR	NR	0.91 (0.323*)	
		LEN 24	Emotional well-being	NR	NR	NR	0.34 (0.319*)	
		LEN 18 vs LEN 24	Emotional well-being	NR	NR	NR	0.57 (95% CI: -0.32-1.46)	0.2076
		LEN18	Functional well-being	NR	NR	NR	-1.56 (0.531*)	
		LEN 24	Functional well-being	NR	NR	NR	-1.28 (0.529*)	
		LEN 18 vs LEN 24	Functional well-being	NR	NR	NR	-0.28 (-1.74-1.19)	0.7076
Hudgens, 2016[94]	FACIT/FACT instruments	First line lenvatinib patients	Overall domain score	NR	NR	74.03	NR	NR

		2nd line lenvatinib patients	Overall domain score	NR	NR	69.92	NR	NR
	EQ-5D-5L index	First line lenvatinib patients	NR	NR	NR	0.76	NR	NR
		2nd line lenvatinib patients	NR	NR	NR	0.71	NR	NR
Raef, 2016[97]	NR	Overall population	NR	NR	NR	NR	"The use of sorafenib is associated with significant AEs and lower QOL score"	NR
Ballal, 2021[98]	VAS	Overall population	VASmax	8.6 (1**)	8 weeks	5.5 (0.8**)	-3.1	0.006

Value is mean (SD) unless otherwise specified. * Value is SE ** Not known if value is SD or SE ***Disutilities
Abbreviations: CI – Confidence interval; EQ-5D – EuroQol 5-dimensional; EORTC-QLQ-C30 – European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire FACIT– The Functional Assessment of Chronic Illness Therapy; FACT – Functional Assessment of Cancer Therapy; LEN – Lenvatinib; NR – Not reported; SD – Standard deviation; SE – Standard error; VAS – Visual analogue scale.
EQ-5D: 0-1 scale (overall); 1 represents the highest possible health state.
EQ-5D-VAS: 0-100 Scale; Higher values indicate better state of health
EORTC-QLQ-C30: 0–100 scale; Higher scores on the functioning scales and on the global health/QoL scale indicate good QoL, while high scores on the symptom scales indicate reduced QoL.
FACIT/FACT: Higher scores for the scales and subscales indicate better quality of life
VAS (pain): 0-100 scale; 0mm indicating no pain and 100mm indicating the worst possible pain)

B.3.4.3 Adverse reactions

Please see Section B.2.10 for the full details of adverse event data in the COSMIC-311 trial.

As standard practice in CEMs and aligning with TA535[76], only TEAEs of grade 3 and above with an incidence of greater than 5% are included. This assumes that these TEAEs included in the CEM are expected to have an impact on healthcare resource use, costs or an impact on HRQoL.

The rates of AEs for cabozantinib and BSC were obtained from the COSMIC-311 trial data. Details regarding sources of AE management costs are provided in Table 38.

Table 30: Treatment-related Adverse Events and Incidence Rates

Parameter	Cabozantinib %	BSC %
Hand–foot syndrome	10	0
Proteinuria	1	0
Hypertension	9	3
Diarrhoea	7	0
Fatigue	8	0
Hypocalcaemia	7	2
Reference	COSMIC-311[46]	

Abbreviations: BSC – Best supportive care

Disutility associated with AEs was included in the base case to assess the impact of disutility associated with AEs. The disutility associated with particular AEs were extracted from TA535[76] and TA498[100] and are shown in Table 31. These disutility values were applied as a one-off decrement upon the health state utility in the first month of the model, under the assumption that AEs were likely to occur very soon after treatment and only require acute care. This approach to modelling AEs is consistent with approaches accepted in previous NICE appraisals[76].

Table 31: Disutility Associated with AEs

AE	Disutility	Reference
Hand-foot syndrome	0.34	TA535[76]
Proteinuria	0	Assumption
Hypertension	0.13	TA498[100]
Diarrhoea	0.47	TA535[76]
Fatigue	0.08	TA535[76]
Hypocalcaemia	0	Assumption

Abbreviations: AE – Adverse event

B.3.4.4 Health-related quality-of-life data used in the cost-effectiveness analysis

In the model, QALYs are used to compare health outcomes, in terms of length and HRQoL across treatment options. QALYs are derived by multiplying the time spent in a specific health state by the health-related utility value associated with that health state.

In the base case, the utility values from Fordham et al. 2015[82] were used based on acceptance in several NICE appraisals, including TA742[77] and TA516[93], and the lack of validity of the COSMIC-311 HRQoL data as described in Section B.3.4.1. Table 32 presents the utility values used in the base case.

Additionally, age-related utility decrements were applied in the model to incorporate the natural decline in QoL associated with increasing age. This was implemented in the model using the regression equation published by Ara and Brazier et al. 2010.[84]

Table 32: Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (standard error)	95% confidence interval	Reference in submission (section and page number)	Justification
Health state utility				
PFS	Fordham et al. 2015[82]: 0.87 (0.19)	N/A	B.3.4.1 Health-related quality-of-life data used in the cost-effectiveness analysis. Page 105	Clinicians validated[77] that Fordham utilities[82] were a better representation of the PFS and PD health states. The limited impact of progression in the COSMIC-311 data was likely a result of limited follow-up in the PD state, as the data suggests that utility falls over time in the PD state. Fordham et al. 2015[82] has been used and accepted in several NICE appraisals, including TA742[77] and TA516.[93]
PD	Fordham et al. 2015[82]: 0.52 (0.28)			
Age-adjusted utilities	**Base case:** Included **Scenario analysis:** Excluded			
Adverse events				
Adverse events	**Base case:** TA535[76], TA498[100], Fordham et al. 2015[82] **Scenario analysis:** Excluded		B.3.4.3 Health-related quality-of-life data used in the cost-	Quantify the impact of AEs on HRQoL.

		effectiveness analysis. Page 117	Applied in the first month of the model under the assumption that AEs were likely to occur very soon after treatment and only require acute care.

Abbreviations: AEs – Adverse events; HRQoL – Health-related quality-of-life; N/A – Not available; PD – Progressed disease; PFS – Progression free survival

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

An economic SLR was conducted on the 14th October 2021 to identify existing studies reported cost and resource use data in the management of adults with RAI-refractory DTC.

Please see Appendices G and I for the methods used to identify all relevant studies, and description of the cost and resource use studies identified.

The review question and sub-questions evaluated in the cost and resource use SLR were:

- What resource use and cost are associated with treatment for RAI-refractory DTC?
  - Which studies investigate the resource use and costs associated with RAI-refractory DTC?
  - What is the design of these studies?
  - How much resource use and costs are associated with drug use, hospitalisation, outpatient visits, adverse events, workdays missed, productivity loss, disease and caregiver burden?

Of the two publications identified to provide insights in the healthcare resource use and costs for RAI-refractory DTC, as shown in Table 33, Gianoukakis et al.[101], reported direct medical resource use including GP/office/clinic visits and inpatient stay, while Abouzaid, 2015[102] reported not only on direct medical resource use but also on total annual healthcare costs associated with treatment of RAI-refractory DTC. Gianoukakis et al.[101] focused especially on the population of interest, as it compared first line RAI-refractory DTC to second line and beyond RAI-refractory DTC patients. While the absolute figures were not provided, they reported that patients on second line and beyond treatment were in need of more care (e.g., doctor visits) compared to patients receiving first line treatment.

Table 33: Summary of healthcare resource use and cost studies

First Author, Year	Country, Cost Year, Currency	Population/ Intervention & Comparator	Type Of Outcome	List Of Outcomes Reported Per Type of Outcomes
Gianoukakis, 2016[101]	US & EU5, NA, NA	Physicians who were treating RAI refractory DTC patients (n=623)	Direct resource use	GP/office/clinic visits (Times seen patients in the last 12 months, Total doctor visits (current treating and other physicians) in the past 12 months); Inpatient stay, ICU and hospital (Number of times hospitalised for DTC disease-associated complications only in the past 12 months, Number of times)
Abouzaid, 2015[102]	NR, NR, USD	RAI refractory, progressive DTC patients	Direct resource use	Outpatient visits (number of times, all cause and thyroid related); Emergency room visits (number of visits, all cause and thyroid related); Inpatient (hospital, ICU) stay (rate of hospitalisation, all cause and thyroid related); In patient (hospital, ICU) stay (length of stay, all cause and thyroid related); Drug cost and use (number of prescriptions, all cause and thyroid related)
			Healthcare costs	Total annual healthcare costs (all cause and thyroid related)

Abbreviations: DTC – Differentiated thyroid cancer; EU – European Union; GP – General Practitioner; NA – Not applicable; NR – Not reported; RAI – Radioiodine; USD – United States dollars

Costs included in the model

As the perspective of this CEA is the NHS and PSS, the NHS reference costs 2020/21[85] was deemed an appropriate source for the cost inputs for healthcare resource use. A targeted literature review was performed to identify adverse events management costs to apply to adverse events rates from COSMIC-311. Treatment costs were sourced from the British National Formulary via the NICE website.[15] Specifically, the following cost components were considered in the model:

- Treatment costs (treatment acquisition and administration)

- Health state costs (monitoring and end-of-life)
- AE costs

No subsequent treatment costs are included since BSC is the only follow-on treatment available for the current patient population. Where necessary costs were inflated to the 2021 cost year using inflation indices annual percentage increase for adult services published by PSSRU.[87]

B.3.5.1 Intervention and comparators' costs and resource use

B.3.5.1.1. Treatment acquisition costs

Treatment acquisition costs were estimated using data on treatment prices, compliance and the dosing schedule. Information on the compliance and dosing schedule were obtained from COSMIC-311 and the SmPC[14] for cabozantinib. Patients on BSC do not receive any active treatment regimen; thus, no drug acquisition costs were incurred by these patients. Costs and presentation of cabozantinib were extracted from the BNF (accessed in August 2022).[15]

Cost per dose was multiplied by the number of doses per month to estimate the cost per month for each treatment. Table 34 depicts the cost and presentation of cabozantinib.

Table 34: Input Related to Treatment Acquisition Costs of Cabozantinib

Treatment	Presentation	List Price (per 30 tablet pack)	Reference
Cabozantinib	Tablet (30 per pack)	20 mg - £5,143.00 40 mg - £5,143.00 60 mg - £5,143.00	BNF[15]

Abbreviations: BNF – British National Formulary

The cost per month (assuming 30.44 administration) was calculated as £5,218.00 at list price.

The cost per month (assuming 30.44 administration) at [redacted] was calculated as [redacted].

Compliance from COSMIC-311 was applied to the total cost per cycle for cabozantinib to reflect the dose patients are expected to receive in clinical practice as a result of dose interruptions, due to AEs or due to missed doses. The compliance rate for cabozantinib in COSMIC-311 was [redacted].[46]

B.3.5.1.2. Administration

Administration costs were based on NHS References costs 2021/22[85] and PSSRU 2021[87]. In cycle 1, a cost of £245 is incurred for SB11Z "Deliver Exclusively Oral Chemotherapy" [85]. In cycle 2+, a cost of £27.00 is incurred, assuming 30 minutes of pharmacist time[87].

The duration of treatment with cabozantinib for which acquisition and administration costs were applied in the model was based on TTD from COSMIC-311 as described in Section B.3.3.4.

The total cost per month for cabozantinib is found in Table 35. At list price, the cost per cycle 1 and cycle 2+ is £5,463.00 and £5,425.00, respectively. At PAS price, the cost per cycle 1 and cycle 2+ is [redacted] and [redacted], respectively.

Table 35: Cabozantinib total cost per month

Treatment cost	List price (£)	PAS price (£)
Acquisition cost	5,218.00	[redacted]
Administration cost cycle 1	245.00	245.00
Administration cost cycle 2+	27.00	27.00
Total cost per cycle 1	5,463.00	[redacted]
Total cost per cycle 2+	5,245.00	[redacted]

Abbreviations: PAS – Patient access scheme.

B.3.5.2 Health-state unit costs and resource use

B.3.5.2.1. Monitoring costs

The resources used and the frequency of use for patient monitoring varied according to the patient's progression status. These inputs were obtained from TA742[77], and cost for each unit resource was collected from NHS Reference Costs 2020/2021[85] (Table 36). The sum product of the resource use and unit costs was calculated to derive the total monitoring cost per monthly cycle by health state; £381.96 for progression-free cabozantinib, £354.88 for progression-free BSC and £268.86 for progressed disease. These costs were applied every cycle in the model for the period of time patients spent in these health states.

Table 36: Input related to healthcare resource utilisation

Cost category	Resource frequency in PFS (per month)	Resource frequency in PD (per month)	Unit cost (£)	Reference
Blood test routine U&Es	1	0.5	1.85	Directly accessed pathology services. Clinical Biochemistry. DAPS04)[85]
Haematology/ Coagulation test	1	0.5	3.63	Directly accessed pathology services. Clinical Biochemistry. DAPS05)[85]
Blood test calcium and magnesium	1	0.5	1.85	Directly accessed pathology services. Clinical Biochemistry. DAPS04)[85]
Liver function test	1	0.5	1.85	
Thyroid function test	1	0.5	1.85	
Consultant led outpatient visits	1	0.5	224.55	Consultant-led, non-admitted face to face attendance, follow up (Medical Oncology - 370/WF01A)[85]
Nurse-led outpatient visits	0.33	0.5	190.59	Non-Consultant-led, non-admitted face to face attendance, follow up (Medical Oncology - 370/WF01A)[85]
CT scan	0.33	0.33	167.31	Computerised Tomography Scan of more than Three Areas (RD27Z)[85]
ECG	Cabozantinib:0.17 BSC: 0.00	0	162.46	Outpatient procedures. Medical procedures (EY51Z)[85]
Total cost per health state		PFS (Cabozantinib)	381.96	
		PFS (BSC)	354.88	
		PD	268.86	

Abbreviation: CT – Computerised tomograph; NHS – National Health Service; PFS – Progression free survival; PD – Progressive disease

B.3.5.2.2. End of life costs

Healthcare costs substantially increase at the end of life due to the high number of hospital and physician visits, especially for cancer patients. End of life costs (Table 37) were calculated in line with those reported in TA535[76]. Cost categories and costs were taken from the 2014 Nuffield Trust research report "Exploring the cost of care at the end of life" and inflated to 2021 GBP based on the latest UK Consumer Price Index data at a value of £8,705.50.[86,103] This cost is applied in the model as a one-off cost when a patients enters the death health state.

Table 37: End of life costs

	Summary costs associated with cancer diagnosis (£)	Inflated costs (£)	Reference
GP visits per person	365	449.20	Georghiou et al. 2014[86]
District nurse per person	588	723.60	
Local authority-funded social care per person	444	505.30	
Emergency inpatient admissions	4,071	4,864.10	
Non-emergency inpatient admissions	1,360	1625.00	
Outpatient attendances	378	451.60	
A&E visits	80	95.60	
Total	7,286	8,705.50	

Abbreviations: A&E – Accident & Emergency; GP – General Practitioner

B.3.5.3 Adverse reaction unit costs and resource use

In order to ascertain costs for managing AEs, resources used to manage an AE were identified with associated costs based on the NHS Schedule of Reference Costs 2020/2021 and PSSRU 2021.[85,87,103] In the model, AE management costs were considered as one-off costs (at first month), estimated as the sum product of the AE incidence (Table 30) and the costs associated with management of each AE. It was assumed that AEs were likely to occur very soon after treatment and only require acute care. This approach to modelling AEs is consistent with approaches accepted in previous NICE appraisals.[55] Table 38 presents the base-case management costs associated with each AE.

Table 38: Costs of managing AEs

Adverse Event	Management Cost	Cabozantinib		BSC		Reference
		Probability per month	Total cost per month	Probability per month	Total cost per month	
Hand–foot syndrome	£490.67	10.00%	£49.07	0.00%	£0.00	JD07K. Skin Disorders without Interventions, with CC Score 0-1. Non-elective Short Stay[85]
Proteinuria	£224.55	1.00%	£2.25	0.00%	£0.00	Consultant-led, non-admitted face to face attendance, follow up (Medical Oncology - 370/WF01A)[85]
Hypertension	£537.86	9.00%	£48.41	3.00%	£16.14	EB04Z Hypertension. Non-elective Short Stay.[85]
Diarrhoea	£635.99	7.00%	£44.52	0.00%	£0.00	FD10M Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2. Non-elective Short Stay[85]
Fatigue	£44.00	8.00%	£3.52	0.00%	£0.00	PSSRU. Community based staff. Nurse unit cost[87]
Hypocalcaemia	£625.96	7.00%	£43.82	2.00%	£12.52	Other Red Blood Cell Disorders with CC Score 0-1. Non-elective short stay[85]
Total			**£191.58**		**£28.66**	

Abbreviations: PSSRU – Personal Social Services Research Unit

B.3.5.4 Miscellaneous unit costs and resource use

No additional costs or resource use were used to inform this cost-effectiveness analysis.

B.3.6. Severity

Ipsen investigated whether locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed during or after prior systemic therapy qualifies for the criteria to be classed as a severe disease. For this population of interest the standard of care is BSC. Therefore, the lifetime QALY gain of patients receiving BSC (as estimated by the CEM) is expressed as a proportion of the estimated lifetime QALY gain of healthy patients of the same age and gender distribution, to understand the extent to which the disease deprives the patient of their remaining QALYs.

The baseline characteristics were based on the COSMIC-311 trial (see Table 39). The PFS and OS data are outlined in Section B.3.3. Utility data and scenario analyses are outlined in Section B.3.10.3.

Following QALY shortfall analysis with different utility values, we conclude that locally advanced or metastatic DTC patients, refractory or not eligible to RAI who have progressed during or after prior systemic therapy qualifies for the 1.2 severity modifier. Therefore, the results of the CEM will be assessed against the willingness to pay thresholds of £24,000 to £36,000 per QALY.

Table 39: Summary features of QALY shortfall analysis

Factor	Value (reference to appropriate table or figure in submission)	Reference to section in submission
Sex distribution	47% male	Section B.2.3.3
Starting age	65 years old	Section B.2.3.3

Table 40: Base case summary of health state benefits and utility values for QALY shortfall analysis

State	Utility value: mean (standard error)	Undiscounted life years
PFS	Fordham et al. 2015: 0.87 (0.19)	0.44
PD	Fordham et al. 2015: 0.52 (0.28)	1.88

Abbreviations: PD – Progressed disease; PFS – Progression free survival; QALY – Quality adjusted life year

Table 41: Summary of QALY shortfall analysis

Utility source	Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with BSC	Absolute QALY shortfall	Proportional
Fordham et al. 2015	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC – Best supportive care; QALY – Quality adjusted life year

B.3.7. Uncertainty

Uncertainty may be derived from the small patient population in COSMIC-311 and the cross-over adjustment analyses for OS.

There are small patient numbers informing the clinical observations. The COSMIC-311 population was small, meaning that variation observed in a few patients drives the clinical measures in the economic analysis which may introduce bias if extreme values are observed.

Traditional ITT analysis is not appropriate for the analysis of the COSMIC-311 OS data as it does not account for the possible OS benefit received by placebo patients who switched to cabozantinib and can therefore underestimate the relative efficacy of cabozantinib compared to a true placebo arm that does not include patients receiving subsequent cabozantinib treatment. The RPSFT method was used, however, this comes with important assumptions and limitations that need to be acknowledged. A limitation of the RPSFT is the “common treatment effect” assumption, that is, the treatment effect received by switchers must be the same as the treatment effect

received by patients initially randomised to the experimental group (i.e., similar efficacy of treatment whether initiated on or switched to at a later time when the disease prognosis for a patient might have changed). Therefore, if patients who switch are also the progressors (i.e., the switch happened after progression, which likely changed the disease prognosis), then it is unlikely that the treatment effect is the same for those initially randomised to treatment and those that switched. Additionally, it should be noted that the placebo patients who switched treatment receive cabozantinib therapy; therefore, in the RPSFT the common treatment effect would implicitly assume that the treatment effect is the same for pre-progression cabozantinib and post-progression cabozantinib patients.

For the Full ITT population (CCO2) the placebo arm included 40 subjects who crossed over to receive cabozantinib. At CCO1 the placebo arm included 19 subjects who subsequently crossed over to receive cabozantinib. For CCO2 and CCO1, the stratified HRs were 0.76 (95% CI: 0.45, 1.31) and 0.54 (95% CI: 0.27, 1.11), respectively. It could be argued that the HR for OS from COO1 is a better reflection of the survival benefit of cabozantinib as less patients had crossed over.

B.3.8. Summary *of base-case analysis inputs and assumptions*

B.3.8.1 Summary of base-case analysis inputs

A summary of variables applied in the economic model is presented in Table 42.

Table 42: Summary of variables applied in the economic model

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Settings			
Time horizon	35	N/A	Section B.3.3
Age at baseline (years)	65	62,68 (Gamma)	Section B.3.3
Percentage male at baseline	47%	(38%, 56%) (Beta)	Section B.3.3
Discount rate costs	3.5%	N/A	Section B.3.2.2

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Discount rate outcomes	3.5%	N/A	Section B.3.2.2
Clinical inputs			
PFS (cabozantinib)	Weibull	N/A	Section B.3.3
PFS (BSC)		N/A	Section B.3.3
OS (cabozantinib)	Exponential	N/A	Section B.3.3
OS (BSC)		N/A	Section B.3.3
TTD (cabozantinib)	Exponential	N/A	Section B.3.3
TTD (BSC)	N/A	N/A	Section B.3.3
Cost inputs			
Cabozantinib acquisition cost list price	£5,143.00	N/A	Section B.3.5.1
Cabozantinib acquisition cost PAS price	[redacted]	N/A	Section B.3.5.1
Cabozantinib administration cost cycle 1	£245.00	£245, £245 (Gamma)	Section B.3.5.1
Cabozantinib administration cost cycle 2+	£27.00	£27, £27 (Gamma)	Section B.3.5.1
Cabozantinib compliance	[redacted]	[redacted] (Beta)	Section B.3.5.1
Blood test routine U&Es	£1.85	N/A	Section B.3.5.2
Haematology/Coagulation test	£3.63	N/A	Section B.3.5.2
Blood test calcium and magnesium	£1.85	N/A	Section B.3.5.2
Liver function test	£1.85	N/A	Section B.3.5.2
Thyroid function test	£1.85	N/A	Section B.3.5.2
Consultant led outpatient visits	£224.55	N/A	Section B.3.5.2
Nurse-led outpatient visits	£190.59	N/A	Section B.3.5.2
CT scan	£167.31	N/A	Section B.3.5.2
ECG	£162.46	N/A	Section B.3.5.2
Hand foot syndrome	£490.67	N/A	Section B.3.5.3
Proteinuria	£224.55	N/A	Section B.3.5.3
Hypertension	£537.86	N/A	Section B.3.5.3
Diarrhoea	£635.99	N/A	Section B.3.5.3
Fatigue	£44.00	N/A	Section B.3.5.3

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Hypocalcaemia	£625.96	N/A	Section B.3.5.3
End of life	£8,705.49	N/A	
Resource use: PFS			
Blood test routine U&Es	1.00	N/A	Section B.3.5.2
Haematology/Coagulation test	1.00	N/A	Section B.3.5.2
Blood test calcium and magnesium	1.00	N/A	Section B.3.5.2
Liver function test	1.00	N/A	Section B.3.5.2
Thyroid function test	1.00	N/A	Section B.3.5.2
Consultant led outpatient visits	1.00	N/A	Section B.3.5.2
Nurse-led outpatient visits	0.33	N/A	Section B.3.5.2
CT scan	0.33	N/A	Section B.3.5.2
ECG (Cabozantinib)	0.17	N/A	Section B.3.5.2
ECG (BSC)	0.00	N/A	Section B.3.5.2
Resource use: PD			
Blood test routine U&Es	0.50	N/A	Section B.3.5.2
Haematology/Coagulation test	0.50	N/A	Section B.3.5.2
Blood test calcium and magnesium	0.50	N/A	Section B.3.5.2
Liver function test	0.50	N/A	Section B.3.5.2
Thyroid function test	0.50	N/A	Section B.3.5.2
Consultant led outpatient visits	0.50	N/A	Section B.3.5.2
Nurse-led outpatient visits	0.50	N/A	Section B.3.5.2
CT scan	0.33	N/A	Section B.3.5.2
ECG	0.00	N/A	Section B.3.5.2
Adverse events probability per month (Cabozantinib)			
Hand foot syndrome	10.00%	N/A	Section B.3.5.3
Proteinuria	1.00%	N/A	Section B.3.5.3
Hypertension	9.00%	N/A	Section B.3.5.3
Diarrhoea	7.00%	N/A	Section B.3.5.3
Fatigue	8.00%	N/A	Section B.3.5.3
Hypocalcaemia	7.00%	N/A	Section B.3.5.3
Adverse events probability per month (BSC)			

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Hand foot syndrome	0.00%	N/A	Section B.3.5.3
Proteinuria	0.00%	N/A	Section B.3.5.3
Hypertension	3.00%	N/A	Section B.3.5.3
Diarrhoea	0.00%	N/A	Section B.3.5.3
Fatigue	0.00%	N/A	Section B.3.5.3
Hypocalcaemia	2.00%	N/A	Section B.3.5.3
QoL inputs			
PFS	0.870	0.84,0.91 (Beta)	Section B.3.4.4
PD	0.520	0.43,0.62 (Beta)	Section B.3.4.4
Hand foot syndrome disutility	0.34	N/A	Section B.3.4.3
Proteinuria disutility	0.00	N/A	Section B.3.4.3
Hypertension disutility	0.13	N/A	Section B.3.4.3
Diarrhoea disutility	0.47	N/A	Section B.3.4.3
Fatigue disutility	0.08	N/A	Section B.3.4.3
Hypocalcaemia disutility	0.00	N/A	Section B.3.4.3

Abbreviations: CT – Computerised tomography; ECG – Electrocardiogram; OS – Overall survival; PD – Progressed disease; PFS – Progression free survival; QOL – Quality of life; TTD – Time to treatment discontinuation

B.3.8.2 Assumptions

Table 43: Assumptions list

Variable	Assumed value	Justification
Time horizon Section B.3.2.2	35 years	Patients entering the model have a mean age of 65 years based on the COSMIC-311 trial baseline characteristics. Patients in the cohort are not expected to live beyond 100 years and therefore a 35-year time horizon was deemed appropriate (100-65 = 35).
Population entering the economic model Section B.3.2.1	Adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed	The population is in line with the EMA and MHRA approval for cabozantinib[1,14] and is the same as the ITT population of the COSMIC-311 phase-3 clinical trial.[46,47]

	during or after prior systemic therapy	
Health states Section B.3.2.1	Progression free, progressed disease and death	Health states are consistent with the natural disease progression in oncology patients.[80]
Clinical inputs		
PFS distribution Section B.3.3.2	Weibull	Based on the goodness of fit (AIC and BIC), visual inspection and clinical opinion[50], it was concluded that Weibull was best-fitting curve and applied to the cabozantinib and BSC arms.
OS distribution Section B.3.3.3	Cross-over (RPFST) adjusted Exponential	To account for the cross-over and obtain an unbiased estimate of the OS benefit associated with cabozantinib, the B.3.2.1 adjustment method was used, in line with NICE DSU TSD 16[81], to adjust for cross-over and estimate the OS associated with BSC. UK clinicians[50] advised the survival estimates for the extrapolations for BSC at 5 and 10 years were overestimated and that 0% of patients would be expected to be alive at 5 years. Of the plausible Exponential and Lognormal curves, the Exponential has a sharper decline in the proportion of patients alive over time and was deemed most appropriate with a cap for BSC applied at five years. Therefore, the Exponential was used in the base to model the OS of patients receiving cabozantinib and BSC with 0% of BSC patients modelled from five years.
TTD distribution Section B.3.3.4	Exponential	Based on the goodness of fit (AIC and BIC), visual inspection the Exponential curve was the best fitting model for the cabozantinib arm.
Cost inputs		
Administration Section B.3.5.1	Cycle 1: £245 Cycle 2+: £27	As per recommendation from UK clinicians at advisory board,[50] administration costs were based on NHS References costs 2021/22[85] and PSSRU 2021[87]. In cycle 1, a

		cost of £245 is incurred for SB11Z "Deliver Exclusively Oral Chemotherapy" [85]. In cycle 2+, a cost of £27.00 is incurred, assuming 30 minutes of pharmacist time[87].
Compliance Section B.3.5.1	Included	Compliance from COSMIC-311 was applied to the total cost per cycle for cabozantinib to reflect the dose patients are expected to receive in clinical practice as a result of dose interruptions due to AEs or due to missed doses.
Adverse event costs Section B.3.5.3	Included	As standard practice in CEMs and aligning with TA535[76], TEAEs of grade 3 and above with an incidence of greater than 5% are included. These were applied as a one-off decrement upon the health state cost in the first month of the model, under the assumption that AEs were likely to occur very soon after treatment and only require acute care.
HRQoL inputs		
Health state utilities Section B.3.4.4	Fordham et al. 2015[82]	Clinicians validated[50] that Fordham utilities[82] were a better representation of the PFS and PD health states. The limited impact of progression in the COSMIC-311 data was likely a result of limited follow-up in the PD state, as the data suggests that utility falls over time in the PD state. In addition, Fordham et al. 2015[82] has been used and accepted in several NICE appraisals, including TA742[77] and TA516.[93]
Adverse event disutilities Section B.3.4.3	Included	As standard practice in CEMs as the health states utilities in Fordham et al. does not account for in trial adverse events and was previously done in TA535[76]. These disutility values were applied as a one-off decrement upon the health state utility in the first month of the model, under the assumption that AEs were likely to occur very soon after treatment and only require acute care.

Abbreviations: AE – Adverse event; AIC – Akaike's Information Criterion; BIC – Bayesian Information Criterion; BSC – Best supportive care; CEM – Cost-effectiveness model; DSU – Decision Support Unit; DTC – Differentiated thyroid cancer; EMA – European Medicines Agency; HRQoL – Health-related quality of life; ITT – Intention to treat; MHRA – Medicines and Healthcare products Regulatory Agency; OS – Overall survival; PD – Progressed disease; PFS – Progression free survival; QOL – Quality of life; RAI – Radioactive iodine; RPFST – Rank preserving structural failure time; TEAE – Treatment emergent adverse event; TSD – Technical Support Document; TTD – Time to treatment discontinuation.

B.3.9. Base-case results

B.3.9.1 Base-case incremental cost-effectiveness analysis results

As described in Section B.1.2, a confidential simple patient access scheme (PAS) has been approved by the Patient Access Schemes Liaison Unit (PASLU). The pack price under this scheme is [redacted] (a [redacted]). This PAS has been applied and the results presented to reflect this discount. The deterministic, base case incremental cost-effectiveness analysis results are presented in Table 44. Cabozantinib was associated with [redacted] incremental costs and [redacted] incremental QALYs compared to BSC, which corresponds to an ICER of £28,148 per QALY gained. Disaggregated base case results are presented in Appendix J.

The net health benefit is displayed in Table 45. The thresholds for net health benefit (NHB) have been updated to align with the willingness to pay threshold outlined in Section B.3.6. The NHB at £36,000 of 0.154 implies that overall population health would be increased as a result of introducing cabozantinib.

Table 44: Deterministic base-case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
PAS price								
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	28,148	28,148

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Table 45: Net health benefit

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	NHB at £24,000	NHB at £36,000
PAS price						
BSC	[redacted]	[redacted]	[redacted]	[redacted]		
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	-0.122	0.154

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; NHB – Net health benefit; QALYs – Quality-adjusted life years

B.3.10. Exploring uncertainty

B.3.10.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analyses (PSA) were conducted to explore the impact of model parameters uncertainty on the results. PSA involves drawing a value at random for each variable from its uncertainty distribution. This is performed for each parameter simultaneously and the resulting incremental results are recorded. This constitutes one 'simulation'. 10,000 simulations were performed, which each gave a distribution of incremental results, and consequently, an assessment of the robustness of the cost-effectiveness results.

For event rates and utilities, a beta distribution was used to restrict draws to between 0 and 1. For costs and resource use estimates, and hazard ratios a gamma distribution was fitted to prevent values less than zero. Treatment costs remained fixed. An incremental cost-effectiveness plane (ICEP) scatter plot (Figure 29), cost-effectiveness acceptability curve (CEAC) (Figure 30) and cost-effectiveness acceptability frontier (CEAF) (Figure 31) were produced to graphically illustrate the level of variability and uncertainty in the results.

The mean values for total costs, LYs, QALYs, and incremental cost per QALY gained for cabozantinib versus BSC for the population of interest generated through 10,000 simulations of the base-case PSA are presented in Table 46. The output shows that on average, cabozantinib results in [redacted] incremental QALYs compared to BSC. In addition, cabozantinib is associated with [redacted] incremental costs over a life-time horizon compared with BSC, resulting in an ICER of £35,249.

Figure 29 to Figure 31 display the ICEP, CEAC and CEAF of cabozantinib versus BSC. The probabilistic results are centred around the deterministic value and the CEAC shows that cabozantinib is cost-effective until £33,000.

Table 46: Probabilistic sensitivity analyses - Base case

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	35,249

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 29: Incremental cost-effectiveness plane– Base case

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 30: Cost-effectiveness acceptability curve - Base case

Abbreviations: BSC – Best supportive care

Figure 31: Cost-effectiveness acceptability frontier - Base case

Abbreviations: BSC – Best supportive care

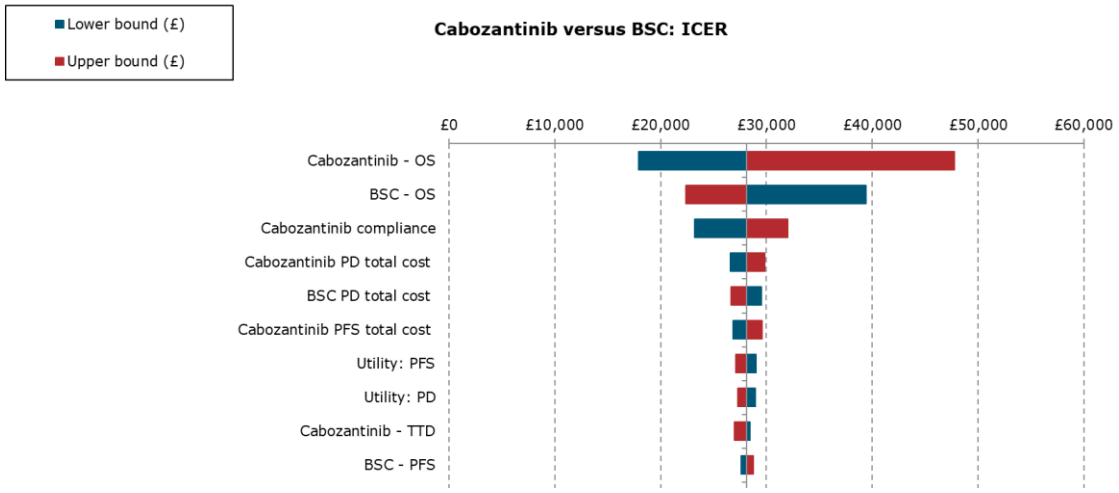
B.3.10.2 Deterministic sensitivity analysis

Deterministic one-way sensitivity analysis (OWSA) was conducted to explore the level of uncertainty in the model results. The OWSA involved varying one parameter at a time and assessing the subsequent impact on the incremental QALYs and incremental costs. By adjusting each parameter individually, the sensitivity of the model results to that parameter can be assessed.

The OWSA was conducted by allocating a 'low' value and a 'high' value to each parameter; the low value is the lower bound of the 95% CI, the high value is the upper bound of the 95% CI. In the absence of CI data, the variable was altered by +/- 10%. A tornado diagram was developed to graphically present the parameters which have the greatest effect on the ICER.

A OWSA tornado diagram presenting the top 10 most sensitive parameters for cabozantinib versus BSC is presented in Figure 32. Table 47 presents the OSWA results for these 10 parameters. The model was most sensitive to the overall survival of cabozantinib and BSC.

Figure 32: One-way sensitivity analysis tornado plot



Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; OS – Overall survival; PD – Progressed disease; PFS – Progression free survival; TTD – Time to treatment discontinuation

Table 47: One-way sensitivity analysis results

Parameter	Lower bound (£)	Upper bound (£)	Difference (£)
Cabozantinib - OS	£17,920	£47,776	£29,857
BSC - OS	£39,416	£22,388	£17,027

Cabozantinib compliance	£23,246	£31,988	£8,743
Cabozantinib PD total cost	£26,581	£29,874	£3,293
BSC PD total cost	£29,519	£26,637	£2,882
Cabozantinib PFS total cost	£26,832	£29,597	£2,766
Utility: PFS	£29,001	£27,085	£1,917
Utility: PD	£28,928	£27,328	£1,600
Cabozantinib - TTD	£28,466	£26,969	£1,498
BSC - PFS	£27,596	£28,777	£1,181

Abbreviations: BSC – Best supportive care; OS – Overall survival; PD – Progressed disease; PFS – Progression free survival; TTD – Time to treatment discontinuation

B.3.10.3 Scenario analysis

Table 48 details deterministic scenario analysis results for cabozantinib versus BSC. Cabozantinib is cost-effective at the £36,000 per QALY threshold (see Section B.3.6) in all scenarios.

Table 48: Deterministic scenario analysis results

Description	Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
Base case	**BSC**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**-**	**-**
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**28,148**	**28,148**
Discount rate: 0%	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	26,165	26,165
Discount rate: 5%	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	28,976	28,976
Age adjusted utilities: excluded	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	27,937	27,937
PFS: Exponential	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	30,567	30,567
PFS: Generalized gamma	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	29,937	29,937
PFS: Gompertz	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	27,848	27,848

Description	Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
PFS: Log logistic	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	27,740	27,740
PFS: Log normal	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	27,718	27,718
OS: Log normal	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,617	19,617

Probabilistic results

In all scenarios (excluding discount rate at 5%) cabozantinib is cost-effective at the £36,000 per QALY threshold (see Section B.3.6).

Discount rate – 0%

Table 49: Probabilistic scenario analysis results - Discount rate 0%

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	32,869

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 33: Incremental cost-effectiveness plane - Discount 0%

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 34: Cost-effectiveness acceptability curve - Discount 0%

Abbreviations: BSC – Best supportive care

Figure 35: Cost-effectiveness acceptability frontier - Discount 0%

Abbreviations: BSC – Best supportive care

Discount rate – 5%

Table 50: Probabilistic scenario analysis results - Discount rate 5%

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	36,332

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 36: Incremental cost-effectiveness plane - Discount 5%

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 37: Cost-effectiveness acceptability curve - Discount 5%

Figure 38: Cost-effectiveness acceptability frontier - Discount 5%

Age-adjusted utilities – excluded

Table 51: Probabilistic scenario analysis results – Age-adjusted utilities excluded

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	26,781

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 39: Incremental cost-effectiveness plane - Age-adjusted utilities excluded

Figure 40: Cost-effectiveness acceptability curve - Age-adjusted utilities excluded

Figure 41: Cost-effectiveness acceptability frontier - Age-adjusted utilities excluded

PFS: Exponential

Table 52: Probabilistic scenario analysis results – PFS exponential

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	35,206

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS –Progression free survival; QALYs – Quality-adjusted life years

Figure 42: Incremental cost-effectiveness plane – PFS exponential

[redacted]

Figure 43: Cost-effectiveness acceptability curve – PFS exponential

Figure 44: Cost-effectiveness acceptability frontier - PFS exponential

PFS: Generalized gamma

Table 53: Probabilistic scenario analysis results – PFS generalized gamma

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	25,224

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS – Progression free survival; QALYs – Quality-adjusted life years

Figure 45: Incremental cost-effectiveness plane – PFS generalized gamma

Figure 46: Cost-effectiveness acceptability curve – PFS generalized gamma

Figure 47: Cost-effectiveness acceptability frontier – PFS generalized gamma

PFS: Gompertz

Table 54: Probabilistic scenario analysis results – PFS gompertz

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	24,967

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS –Progression free survival; QALYs – Quality-adjusted life years

Figure 48: Incremental cost-effectiveness plane – PFS gompertz

[redacted]

Figure 49: Cost-effectiveness acceptability curve – PFS gompertz

Figure 50: Cost-effectiveness acceptability frontier – PFS gompertz

PFS: Log logistic

Table 55: Probabilistic scenario analysis results – PFS log logistic

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	32,931

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS –Progression free survival; QALYs – Quality-adjusted life years

Figure 51: Incremental cost-effectiveness plane – PFS log logistic

[redacted]

Figure 52: Cost-effectiveness acceptability curve – PFS log logistic

Figure 53: Cost-effectiveness acceptability frontier – PFS log logistic

PFS: Log normal

Table 56: Probabilistic scenario analysis results – PFS log normal

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	30,415

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS –Progression free survival; QALYs – Quality-adjusted life years

Figure 54: Incremental cost-effectiveness plane – PFS log normal

Figure 55: Cost-effectiveness acceptability curve – PFS log normal

Figure 56: Cost-effectiveness acceptability frontier – PFS log normal

OS: Log normal

Table 57: Probabilistic scenario analysis results – OS log normal

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,094

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; OS – Overall survival; QALYs – Quality-adjusted life years

Figure 57: Incremental cost-effectiveness plane - OS lognormal

Figure 58: Cost-effectiveness acceptability curve - OS lognormal

Figure 59: Cost-effectiveness acceptability frontier - OS lognormal

B.3.11. Subgroup analysis

No subgroup analyses were considered.

B.3.12. Benefits not captured in the QALY calculation

Currently there are no interventions for patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who progressed during or after prior systemic therapy. Availability of a treatment option for patients will create a standardisation that will facilitate follow-on treatment to improve patient lives.

B.3.13. Validation

B.3.13.1 Validation of cost-effectiveness analysis

- Where possible, insights from the NICE submission (TA535) were utilised within the cost-effectiveness model[104].

- An internal validity check was performed by the model developers. This included a quality check of model codes, model inputs including both a comparison to the original source and any intermediate calculations, and a check of model output. The model was developed by two independent health economists.

- PFS extrapolations, OS extrapolations, administration cost and utilities inputs and assumptions were validated by three UK clinicians, all of which have experience treating patients with DTC, who attended an advisory board. The outcome of the advisory board is referenced throughout the B3 section.

B.3.14. Interpretation and conclusions of economic evidence

At PAS price, over a 35-year time horizon, deterministic base-case results demonstrated that cabozantinib accrued [redacted] QALYs at a cost of [redacted], whilst patients receiving BSC accrued [redacted] QALYs at a cost of [redacted]. The resulting ICER in the base case was £28,148 per QALY, well below the NICE threshold of £36,000 per QALY based on the severity modifier calculation.

The PSA output shows that on average, cabozantinib results in [redacted] incremental QALYs compared to BSC. In addition, cabozantinib is associated with [redacted] incremental costs over a life-time horizon compared with BSC, resulting in an ICER of £35,249.

OWSA found that results were most sensitive to the overall survival of cabozantinib and BSC. A variety of deterministic scenario analyses investigating variations in discount rates, utilities, and clinical efficacy all resulted in cabozantinib being cost effective at the £36k per QALY threshold in all scenarios. This is true for probabilistic scenarios also, with the exception of the discount rate at 5% which take the ICER just over the WTP threshold of £36k per QALY.

Overall, the deterministic base-case results, results of the base-case probabilistic sensitivity analysis and all scenario analyses results strongly indicate that cabozantinib is a cost-effective use of NHS resources.

B.4. References

1. MHRA. Summary of Product Characteristics [Internet]. 2022 [cited 2022 Sep 1]. Available from: https://mhraproducts4853.blob.core.windows.net/docs/79d963db71efbd44018aa286cfa797b9df20328b

2. Filetti S, Durante C, Hartl D, Leboulleux S, Locati L, Newbold K, et al. Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2019;30(12):1856–83.

3. Filetti S, Durante C, Hartl D, Leboulleux S, Locati L, Newbold K, et al. ESMO Clinical Practice Guideline update on the use of systemic therapy in advanced thyroid cancer. Ann Oncol. 2022;33(7):674–84.

4. National Institute for Health and Care Excellence. Overview | Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine | Guidance | NICE | TA535 [Internet]. NICE; [cited 2022 Jul 4]. Available from: https://www.nice.org.uk/guidance/ta535

5. NHS England. National Cancer Drugs Fund List ver1.209 07-Apr-22 (Including list of NICE approved and baseline funded drugs/indications from 1st April 2016 with criteria for use) [Internet]. 2022. Available from: https://www.england.nhs.uk/wp-content/uploads/2017/04/national-cdf-list-v1.209.pdf

6. National Institute for Health and Care Excellence. Overview | Selpercatinib for treating advanced thyroid cancer with RET alterations | Guidance | NICE | TA742 [Internet]. NICE; [cited 2022 Jul 4]. Available from: https://www.nice.org.uk/guidance/ta742

7. Overview | Entrectinib for treating NTRK fusion-positive solid tumours | Guidance | NICE | TA644 [Internet]. NICE; [cited 2022 Jul 12]. Available from: https://www.nice.org.uk/guidance/ta644

8. National Institute for Health and Care Excellence. Overview | Larotrectinib for treating NTRK fusion-positive solid tumours | Guidance | NICE | TA630 [Internet]. NICE; [cited 2022 Jul 25]. Available from: https://www.nice.org.uk/guidance/ta630

9. Cancer Research UK. Thyroid cancer incidence statistics [Internet]. 2021 [cited 2022 Jul 12]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/thyroid-cancer/incidence

10. Sennino B, McDonald DM. Controlling escape from angiogenesis inhibitors. Nat Rev Cancer. 2012;12(10):699–709.

11. Harshman LC, Choueiri TK. Targeting the hepatocyte growth factor/c-Met signaling pathway in renal cell carcinoma. Cancer J. 2013;19(4):316–23.

12. Sennino B, Ishiguro-Oonuma T, Wei Y, Naylor RM, Williamson CW, Bhagwandin V, et al. Suppression of tumor invasion and metastasis by concurrent inhibition of c-Met

and VEGF signaling in pancreatic neuroendocrine tumors. Cancer Discov. 2012;2(3):270–87.

13. Yakes FM, Chen J, Tan J, Yamaguchi K, Shi Y, Yu P, et al. Cabozantinib (XL184), a novel MET and VEGFR2 inhibitor, simultaneously suppresses metastasis, angiogenesis, and tumor growth. Mol Cancer Ther. 2011;10(12):2298–308.

14. EMA. Summary of Product Characteristics [Internet]. 2022 [cited 2022 Aug 10]. Available from: https://www.ema.europa.eu/en/documents/product-information/cabometyx-epar-product-information_en.pdf

15. BNF. Medicinal forms | Cabozantinib | Drugs | BNF content published by NICE [Internet]. [cited 2022 Aug 4]. Available from: https://bnf.nice.org.uk/drugs/cabozantinib/medicinal-forms/

16. 2022 ICD-10-CM Diagnosis Code C73: Malignant neoplasm of thyroid gland [Internet]. [cited 2021 Nov 19]. Available from: https://www.icd10data.com/ICD10CM/Codes/C00-D49/C73-C75/C73-/C73

17. Cancer Research UK. Types of thyroid cancer [Internet]. [cited 2022 Jul 12]. Available from: https://www.cancerresearchuk.org/about-cancer/thyroid-cancer/stages-types/types

18. Lirov R, Worden FP, Cohen MS. The treatment of advanced thyroid cancer in the age of novel targeted therapies. Drugs. 2017;77(7):733–45.

19. Kent WD, Hall SF, Isotalo PA, Houlden RL, George RL, Groome PA. Increased incidence of differentiated thyroid carcinoma and detection of subclinical disease. Cmaj. 2007;177(11):1357–61.

20. Tumino D, Frasca F, Newbold K. Updates on the management of advanced, metastatic, and radioiodine refractory differentiated thyroid cancer. Front Endocrinol. 2017;8:312.

21. Rossi ED, Pantanowitz L, Hornick JL. A worldwide journey of thyroid cancer incidence centred on tumour histology. Lancet Diabetes Endocrinol. 2021 Apr;9(4):193–4.

22. Miranda-Filho A, Lortet-Tieulent J, Bray F, Cao B, Franceschi S, Vaccarella S, et al. Thyroid cancer incidence trends by histology in 25 countries: a population-based study. Lancet Diabetes Endocrinol. 2021 Apr;9(4):225–34.

23. Gild et al. ML. Clinical guidance for radioiodine refractory differentiated thyroid cancer. Clin Endocrinol Oxf. 2018;88:529-537,.

24. Xu B, Ghossein R. Poorly differentiated thyroid carcinoma. Semin Diagn Pathol. 2020 Sep;37(5):243–7.

25. International Agency for Research on Cancer. United Kingdom [Internet]. Available from: https://gco.iarc.fr/today/data/factsheets/populations/826-united-kingdom-fact-sheets.pdf?msclkid=05ebee3ea3b911eca463c9ff3df72273

26. Dal Maso L, Tavilla A, Pacini F, Serraino D, van Dijk BAC, Chirlaque MD, et al. Survival of 86,690 patients with thyroid cancer: A population-based study in 29 European countries from EUROCARE-5. Eur J Cancer. 2017 May;77:140–52.

27. Office for National Statistics. Cancer survival in England - adults diagnosed [Internet]. 2019. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed

28. Durante C, Haddy N, Baudin E, Leboulleux S, Hartl D, Travagli JP, et al. Long-term outcome of 444 patients with distant metastases from papillary and follicular thyroid carcinoma: benefits and limits of radioiodine therapy. J Clin Endocrinol Metab. 2006 Aug;91(8):2892–9.

29. Babu G, Kainickal CT. Update on the systemic management of radioactive iodine refractory differentiated thyroid cancer (Review). Mol Clin Oncol. 2021 Feb;14(2):35.

30. Oh HS, Shin DY, Kim M, Park SY, Kim TH, Kim BH, et al. Extended Real-World Observation of Patients Treated with Sorafenib for Radioactive Iodine-Refractory Differentiated Thyroid Carcinoma and Impact of Lenvatinib Salvage Treatment: A Korean Multicenter Study. Thyroid Off J Am Thyroid Assoc. 2019 Dec;29(12):1804–10.

31. Brilli L, Dalmiglio C, Pilli T, Barbato F, Maino F, Capezzone M, et al. Improvement of Overall Survival Using TKIs as Salvage Therapy in Advanced Thyroid Carcinoma: Real-Life Data on a Single Center Experience. J Clin Med. 2021 Jan 20;10(3):384.

32. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid Off J Am Thyroid Assoc. 2016 Jan;26(1):1–133.

33. Fugazzola L, Elisei R, Fuhrer D, Jarzab B, Leboulleux S, Newbold K, et al. 2019 European Thyroid Association Guidelines for the Treatment and Follow-Up of Advanced Radioiodine-Refractory Thyroid Cancer. Eur Thyroid J. 2019 Oct;8(5):227–45.

34. Schmidt A, Iglesias L, Klain M, Pitoia F, Schlumberger MJ. Radioactive iodine-refractory differentiated thyroid cancer: an uncommon but challenging situation. Arch Endocrinol Metab. 2017 Feb;61(1):81–9.

35. Kaae AC, Kreissl MC, Krüger M, Infanger M, Grimm D, Wehland M. Kinase-Inhibitors in Iodine-Refractory Differentiated Thyroid Cancer—Focus on Occurrence, Mechanisms, and Management of Treatment-Related Hypertension. Int J Mol Sci. 2021 Nov 12;22(22):12217.

36. Perros P, Boelaert K, Colley S, Evans C, Evans RM, Gerrard Ba G, et al. Guidelines for the management of thyroid cancer. Clin Endocrinol (Oxf). 2014 Jul;81 Suppl 1:1–122.

37. Haddad RI, Nasr C, Bischoff L, Busaidy NL, Byrd D, Callender G, et al. NCCN Guidelines Insights: Thyroid Carcinoma, Version 2.2018. J Natl Compr Cancer Netw JNCCN. 2018 Dec;16(12):1429–40.

38. Pacini F, Fuhrer D, Elisei R, Handkiewicz-Junak D, Leboulleux S, Luster M, et al. 2022 ETA Consensus Statement: What are the indications for post-surgical radioiodine therapy in differentiated thyroid cancer? Eur Thyroid J. 2022 Jan 1;11(1):e210046.

39. Haddad RI, Bischoff L, Ball D, Bernet V, Blomain E, Busaidy NL, et al. Thyroid Carcinoma, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2022 Aug 1;20(8):925–51.

40. Gild ML, Tsang VHM, Clifton-Bligh RJ, Robinson BG. Multikinase inhibitors in thyroid cancer: timing of targeted therapy. Nat Rev Endocrinol. 2021 Apr;17(4):225–34.

41. Cabanillas ME, Hu MI, Durand JB, Busaidy NL. Challenges associated with tyrosine kinase inhibitor therapy for metastatic thyroid cancer. J Thyroid Res. 2011;2011:985780.

42. Porter A, Wong DJ. Perspectives on the Treatment of Advanced Thyroid Cancer: Approved Therapies, Resistance Mechanisms, and Future Directions. Front Oncol [Internet]. 2021 [cited 2022 Aug 17];10. Available from: https://www.frontiersin.org/articles/10.3389/fonc.2020.592202

43. Amatu A, Sartore-Bianchi A, Siena S. NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. ESMO Open. 2016;1(2):e000023.

44. Rankin EB, Fuh KC, Castellini L, Viswanathan K, Finger EC, Diep AN, et al. Direct regulation of GAS6/AXL signaling by HIF promotes renal metastasis through SRC and MET. Proc Natl Acad Sci U S A. 2014 Sep 16;111(37):13373–8.

45. Gibney GT, Aziz SA, Camp RL, Conrad P, Schwartz BE, Chen CR, et al. c-Met is a prognostic marker and potential therapeutic target in clear cell renal cell carcinoma. Ann Oncol Off J Eur Soc Med Oncol. 2013 Feb;24(2):343–9.

46. Exelixis, Inc. Clinical Study Report XL184-311. 30 April 2020. [Data on file].

47. Brose MS, Robinson B, Sherman SI, Jarzab B, Lin CC, Vaisman F, et al. Cabozantinib versus placebo in patients with radioiodine-refractory differentiated thyroid cancer who have progressed after prior VEGFR-targeted therapy: Results from the phase 3 COSMIC-311 trial. J Clin Oncol. 2021 May 20;39(15_suppl):6001–6001.

48. Brose MS, Robinson B, Sherman SI, Krajewska J, Lin CC, Vaisman F, et al. Cabozantinib for radioiodine-refractory differentiated thyroid cancer (COSMIC-311): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol. 2021 Aug 1;22(8):1126–38.

49. Ipsen. Advisory Board Report 2021. [Data on file]. 2021.

50. Ipsen. IPSEN HTA Advisory Board for Differentiated Thyroid Cancer. [Data on file]. 2022.

51. Exelixis Inc. Clinical Study Report XL184-311 Addendum 1. 21 May 2021. [Data on file].

52. Hessel C, Mangeshkar M, Motzer RJ, Escudier B, Powles TB, Schwab G, et al. Evaluation of the novel "trial within a trial" design of METEOR, a randomized phase 3 trial of cabozantinib versus everolimus in patients (pts) with advanced renal cell carcinoma (RCC). Ann Oncol. 2016 Oct 1;27:vi284.

53. Latimer NR, Abrams KR. NICE DSU Technical Support Document 16: Adjusting Survival Time Estimates in the Presence of Treatment Switching.

54. COSMIC-311 trial treatment crossover adjustment – technical report. [data on file]. 2021.

55. NICE. Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [ID1059] [Internet]. NICE; 2017 [cited 2022 Aug 18]. Available from: https://www.nice.org.uk/guidance/ta535/documents/committee-papers

56. Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. J Clin Epidemiol. 2005 Dec;58(12):1217–9.

57. Brose MS, Robinson BG, Sherman SI, Jarzab B, Lin CC, Vaisman F, et al. Cabozantinib for previously treated radioiodine-refractory differentiated thyroid cancer: Updated results from the phase 3 COSMIC-311 trial. Cancer. 2022 Dec 15;128(24):4203–12.

58. Brose MS, Robinson BG, Sherman SI, Jarzab B, Lin CC, Vaisman F, et al. Cabozantinib for previously treated radioiodine-refractory differentiated thyroid cancer: Updated results from the phase 3 COSMIC-311 trial (Supplementary material). Cancer. 2022 Dec 15;128(24):4203–12.

59. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials.

60. Brose MS, Shenoy S, Bhat N, Harlacker AK, Yurtal RK, Posey ZA, et al. A phase II trial of cabozantinib (CABO) for the treatment of radioiodine (RAI)-refractory differentiated thyroid carcinoma (DTC) in the first-line setting. J Clin Oncol. 2018 May 20;36(15_suppl):6088–6088.

61. Cabanillas MEDS, Geyer JA, Wirth S, Menefee LJ, Liu ME, Shah SV, et al. Cabozantinib as salvage therapy for patients with tyrosine kinase inhibitor-refractory differentiated thyroid cancer: Results of a multicenter phase II international thyroid oncology group trial. J Clin Oncol. 2017;35:3315-3321,.

62. Brose MS, Nutting CM, Jarzab B, Elisei R, Siena S, Bastholt L, et al. Sorafenib in locally advanced or metastatic, radioactive iodine-refractory, differentiated thyroid

cancer: a randomized, double-blind, phase 3 trial. Lancet. 2014 Jul 26;384(9940):319–28.

63. Schlumberger M, Tahara M, Wirth LJ, Robinson B, Brose MS, Elisei R, et al. Lenvatinib versus Placebo in Radioiodine-Refractory Thyroid Cancer. N Engl J Med. 2015 Feb 12;372(7):621–30.

64. Abou-Alfa GK, Meyer T, Cheng AL, El-Khoueiry AB, Rimassa L, Ryoo BY, et al. Cabozantinib in Patients with Advanced and Progressing Hepatocellular Carcinoma. N Engl J Med. 2018 Jul 5;379(1):54–63.

65. Choueiri TK, Escudier B, Powles T, Tannir NM, Mainwaring PN, Rini BI, et al. Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. Lancet Oncol. 2016 Jul 1;17(7):917–27.

66. Exelixis Inc. Clinical Study Report XL184-311 Addendum 2. 19 August 2021. [Data on file].

67. Sorafenib in radioactive iodine-refractory, locally advanced or metastatic differentiated thyroid cancer: a randomised, double-blind, phase 3 trial - PubMed [Internet]. [cited 2022 Aug 17]. Available from: https://pubmed.ncbi.nlm.nih.gov/24768112/

68. Escudier B, Lougheed JC, Albiges L. Cabozantinib for the treatment of renal cell carcinoma. Expert Opin Pharmacother. 2016 Dec;17(18):2499–504.

69. Centre for Reviews and Dissemination. Systematic Reviews [Internet]. 2009. Available from: https://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

70. Carlson J, Suh K, Xia F, Williamson T, Sullivan S. PCN11 The potential long-term comparative effectiveness of larotrectinib vs. lenvatinib or sorafenib for treatment of NTRK fusion-positive metastatic thyroid cancer. Value Health. 2021 Jun 1;24:S20.

71. Huang W, Chen L, Ting J, Cao V, Sung H, Yokokura M, et al. The Evpi of Treatment Strategies for Radioiodine-Refractory Differentiated Thyroid Cancer. Value Health. 2016 Nov 1;19(7):A886.

72. Wilson L, Huang W, Chen L, Ting J, Cao V. Cost Effectiveness of Lenvatinib, Sorafenib and Placebo in Treatment of Radioiodine-Refractory Differentiated Thyroid Cancer. Thyroid. 2017;27:1043–52.

73. Carrasquilla-Sotomayor M, Alvis-Zakzuk N, Gomez de la Rosa F, Alvis Zakzuk J, Marrugo Figueroa R, Miranda Machado P, et al. Cost and Effectiveness of the use of Sorafenib in Differentiated Thyroid Cancer. Value Health. 2017 Oct;20(9):A876.

74. Tremblay G, Pelletier C, Copher R, Forsythe A, Majethia U. Cost-Effectiveness Analysis of Lenvatinib as a Treatment for Radioactive Iodine Refractory Differentiated Thyroid Cancer in the United States. Value Health. 2016 May;19(3):A151.

75. Erdal E, Sayman H, Turkmen C, Aral F, Yildiz O, Okutur K, et al. Cost-Effectiveness of Sorafenib for Treatment of Radioactive Iodine (Rai)-Refractory Locally Advanced/Metastatic Differentiated Thyroid Cancer (Dtc) In Turkey. Value Health. 2015 May;18(3):A203–4.

76. NICE. Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [ID1059] [Internet]. 2018. Available from: https://www.nice.org.uk/guidance/ta535/documents/committee-papers

77. NICE. Selpercatinib for treating advanced thyroid cancer with RET alterations [ID3744] [Internet]. 2021. Available from: https://www.nice.org.uk/guidance/ta742/documents/committee-papers

78. NICE. Larotrectinib for treating NTRK fusionpositive advanced solid tumours [ID1299] [Internet]. 2020. Available from: https://www.nice.org.uk/guidance/ta630/documents/committee-papers

79. NICE Decision Support Unit [Internet]. [cited 2021 Sep 23]. Available from: http://nicedsu.org.uk/

80. Fleeman N, Houten R, Chaplin M, Beale S, Boland A, Dundar Y, et al. A systematic review of lenvatinib and sorafenib for treating progressive, locally advanced or metastatic, differentiated thyroid cancer after treatment with radioactive iodine. BMC Cancer. 2019 Dec 12;19(1):1209.

81. Latimer NR, Abrams KR. NICE DSU Technical Support Document 16: Adjusting Survival Time Estimates In The Presence Of Treatment Switching [Internet]. School of Health And Related Research; 2014. Available from: https://eclass.uoa.gr/modules/document/file.php/MED1174/Lectures/Unit%207%20%28Missing%20data%20and%20non-compliance%29/TSD16_Treatment_Switching.pdf

82. Fordham BA, Kerr C, de Freitas HM, Lloyd AJ, Johnston K, Pelletier CL, et al. Health state utility valuation in radioactive iodine-refractory differentiated thyroid cancer. Patient Prefer Adherence. 2015;9:1561–72.

83. NICE. NICE health technology evaluations: the manual [Internet]. 2022. Available from: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741

84. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health J Int Soc Pharmacoeconomics Outcomes Res. 2010 Aug;13(5):509–18.

85. NHS England. National Cost Collection: National Schedule of NHS costs - Year 2020-21 - NHS trust and NHS foundation trusts [Internet]. 2022. Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#ncc1819

86. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. [Internet]. Nuffield Trust; 2014. Available from: https://www.nuffieldtrust.org.uk/files/2017-01/end-of-life-care-web-final.pdf

87. Jones K, Burns A. Unit Costs of Health and Social Care 2021. 2021.

88. Latimer NR. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. [Internet]. School of Health And Related Research; 2013. Available from: https://www.ncbi.nlm.nih.gov/books/NBK395885/pdf/Bookshelf_NBK395885.pdf

89. Allison A, White IR, Bond S. rpsftm: An R Package for Rank Preserving Structural Failure Time Models. R J. 2017 Dec 4;9(2):342–53.

90. Korhonen P, Zuber E, Branson M, Hollaender N, Yateman N, Katiskalahti T, et al. Correcting overall survival for the impact of crossover via a rank-preserving structural failure time (RPSFT) model in the RECORD-1 trial of everolimus in metastatic renal-cell carcinoma. J Biopharm Stat. 2012;22(6):1258–71.

91. EQ-5D-5L – EQ-5D [Internet]. 2021 [cited 2022 Aug 19]. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/

92. Hernández-Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. J Health Econ. 2017;55(C):139–52.

93. NICE. Cabozantinib and vandetanib for treating unresectable locally advanced or metastatic medullary thyroid cancer [ID56] [Internet]. 2017. Available from: https://www.nice.org.uk/guidance/ta516/documents/committee-papers

94. Hudgens et al. Quality of survival in radioiodinere fractory thyroid cancer patents treated with lenvatinib. Thyroid. 2016;26(A64).

95. Giani C, Valerio L, Bongiovanni A, Durante C, Grani G, Ibrahim T, et al. Safety and Quality-of-Life Data from an Italian Expanded Access Program of Lenvatinib for Treatment of Thyroid Cancer. Thyroid Off J Am Thyroid Assoc. 2021 Feb;31(2):224–32.

96. Taylor MH, Leboulleux S, Panaseykin Y, Konda B, De la Fouchardière C, Hughes B, et al. 1746P Health-related quality-of-life (HRQoL) analyses from study 211: A phase 2 study in patients (pts) with radioiodine-refractory differentiated thyroid cancer (RR-DTC) treated with 2 starting doses of lenvatinib (LEN). Ann Oncol. 2021 Sep 1;32:S1205.

97. Raef et al. Sorafenib efficacy and safety in advanced well differentiated thyroid cancer (WDTC), early local experience in Saudi Arabia. Thyroid. 2016;26(A154).

98. Ballal S, Yadav M, Moon ES, Kumari S, Roesch F, Tripathi M, et al. First clinical experience and initial outcomes of 177Lu-DOTAGA.(SA.FAPi)2 therapy in patients

with end-stage radioiodine-refractory differentiated thyroid cancer: A Salvage treatment option. J Nucl Med. 2021 May 1;62(supplement 1):1701–1701.

99. Kerr C, Fordham B, Freitas HM de, Pelletier CL, Lloyd A. Health State Utility Valuation in Radio-Iodine Refractory Differentiated Thyroid Cancer (RR-DTC). Value Health. 2014;7(17):A646.

100. NICE. Lenvatinib with everolimus for previously treated advanced renal cell carcinoma [ID1029] [Internet]. 2017. Available from: https://www.nice.org.uk/guidance/ta498/documents/committee-papers

101. Gianoukakis AG, Flores NM, Pelletier CL, Forsythe A, Wolfe GR, Taylor MH. Treatment patterns, health state, and health care resource utilization of patients with radioactive iodine refractory differentiated thyroid cancer. Cancer Manag Res. 2016;8:67–76.

102. Abouzaid S, Li X, Rietschel P. Health Care Utilization And Cost Associated With Radioiodine Refractory (Rai-R) Differentiated Thyroid Cancer (DTC). Value Health. 2015;3(18):A198.

103. Jones K, Burns A. Sources of information. Unit Costs of Health and Social Care 2021 [Internet]. Personal Social Services Research Unit; 2021. Available from: https://www.pssru.ac.uk/pub/uc/uc2021/sourcesofinformation.pdf

104. NICE. Ravulizumab for treating paroxysmal nocturnal haemoglobinuria [Internet]. National Institute for Health and Care Excellence; 2021. Available from: https://www.nice.org.uk/guidance/gid-ta10690/documents/final-appraisal-determination-document

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Summary of Information for Patients (SIP)

September 2022

File name	Version	Contains confidential information	Date
ID4046_Cabozantinib for DTC_SIP_Final	V1.0	No	16/09/2022

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the Health Technology Assessment International – Patient & Citizens Involvement Group (HTAi PCIG). Information about the development is available in an open-access IJTAHC journal article

SECTION 1: Submission summary

1a) Name of the medicine (generic and brand name):

Response: Generic name - cabozantinib (brand name - Cabometyx®)

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

Response: Cabozantinib is used to treat locally advanced or metastatic differentiated thyroid cancer (DTC), a type of cancer in the thyroid gland, in adults when radioactive iodine (RAI) and anticancer medicine treatments are no longer stopping the disease from progressing.

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

Response: An application for the marketing authorisation for cabozantinib in this indication was submitted to the European Medicines Agency (EMA) on 27th July 2021, with the European Centralised decision (considered as final approval) received on 29th April 2022. Cabometyx | European Medicines Agency (europa.eu) The European Commission (EC) decision was provided to the Medicines and Healthcare products Regulatory Agency (MHRA) to facilitate the recognition route, using the EMA approval. Ipsen received GB approval for the Type II extension of the indication in DTC for Cabometyx dated 10th May 2022 from the MHRA.

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

Response: British Thyroid Foundation participated in a series of educational seminars run by Ipsen, in partnership with 14 patient organisations, spanning a wide and diverse range of disease areas. The aim of the three-seminar series was to help the patient groups involved develop a deeper understanding of bench-to-bedside medicine development and the Health Technology Assessment (HTA) process.

No conflict of interest with any of the patient groups and no financial support provided.

SECTION 2: Current landscape

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England.

Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.

Response: Thyroid cancer is a rare type of cancer that affects the thyroid gland, accounting for 1.2% of all new cases of cancer in the UK in 2020. There was a 5-year prevalence of 21,306 people with thyroid cancer in the UK in 2020[1]. Females are much more likely to be diagnosed with thyroid cancer, making up 72% of thyroid cases in the UK[2]. Thyroid cancers can be differentiated or undifferentiated, with differentiated thyroid cancer cells retaining the appearance of normal thyroid cells and usually growing more slowly. Two common types of DTC (papillary and follicular cancers) have similar management and prognosis. Differentiated thyroid cancers are the most common types of thyroid cancers, with papillary cancers responsible for 90% of cases[3]. Typically, DTC is curable, with 10-year survival typically around 85%[4]. The survival of patients with DTC is strongly related to the stage of disease, as once the cancer has spread beyond the thyroid (metastasised), survival rates decrease[5].

DTC usually has a good prognosis when treated with surgery, thyroid-stimulating hormone suppression or RAI, used to destroy remaining cancer cells[6-8]. In recent years targeted therapies have started to be used for residual or recurrent disease. These therapies (tyrosine kinase inhibitors [TKIs]) inhibit signalling pathways that are enhanced during cancer growth[9]. The survival rates of patients with DTC after receiving RAI therapy which has not worked or is no longer stopping the disease from progressing (RAI-refractory) are significantly lower, with the 5-year, 10-year and 15-year survival rates being 66%, 10% and 6% respectively. Mortality rates become much worse for patients following progression from first-line therapy (lenvatinib or sorafenib) if no additional therapy is received[10-12]. The last decade has seen substantial research and development into novel targeted agents to treat patients with RAI-refractory DTC, however there is still an unmet need for patients to have a treatment option after first-line therapy has failed[13-14].

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?

Response: The diagnosis of DTC begins with the neck of the patient being examined for a lump, before further tests are conducted or a possible referral to a specialist if required. A blood test, a thyroid function test, will detect if there are abnormal levels of thyroid hormones in the patients'

blood, however this does not indicate that the patient may have thyroid cancer. An ultrasound scan will be needed if thyroid hormone levels are normal, to create an image of the patients' neck so the doctor can check of a lump that might be cancerous. If a potentially cancerous lump is found, a biopsy can be performed to confirm diagnosis. The only way to confirm if a lump is malignant is to take a biopsy, normally done as an outpatient procedure. Further tests may need to be conducted if a biopsy finds thyroid cancer, as it may be needed to see if the cancer has spread to other regions of the body; these tests are usually a CT scan or an MRI scan[15]. It is recommended to perform an electrocardiogram (ECG) test and liver function tests before cabozantinib treatment and to monitor these during treatment. Thyroid function should be monitored periodically throughout treatment with cabozantinib[16].

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

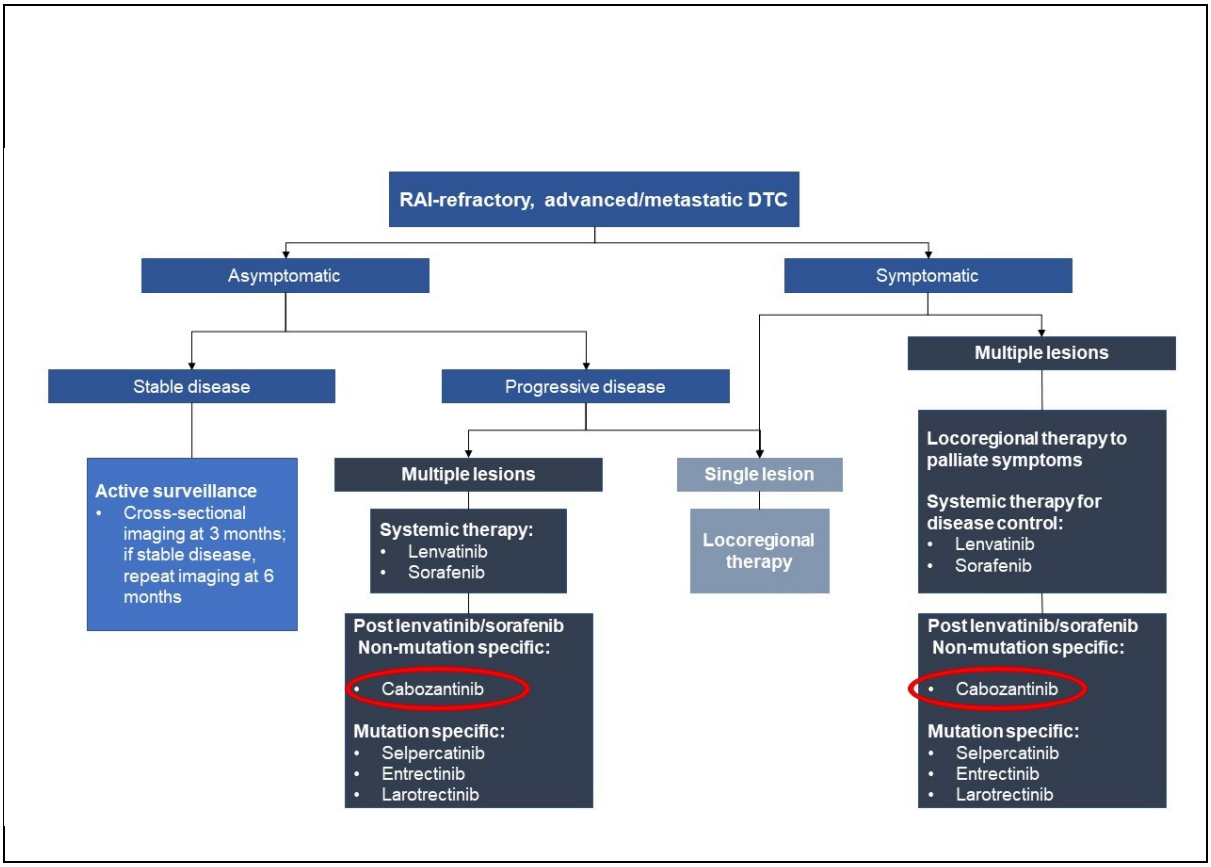
- What is the treatment pathway for this condition and where in this pathway the medicine is likely to be used? Please use diagrams to accompany text where possible. Please give emphasis to the specific setting and condition being considered by NICE in this review. For example, by referencing current treatment guidelines. It may be relevant to show the treatments people may have before and after the treatment under consideration in this SIP.
- Please also consider:
    - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
    - are there any drug–drug interactions and/or contraindications that commonly cause challenges for patient populations? If so, please explain what these are.

Response: NICE Technology Appraisal number 535 (TA535) recommends lenvatinib and sorafenib as the standard first-line therapy for RAI-refractory DTC, with other mutation-specific therapies also recommended following lenvatinib and sorafenib[17]. However, there is limited guidance on second-line and subsequent treatments for non-mutation specific RAI-refractory DTC, with currently only best supportive care (BSC) (e.g., treatments to manage symptoms of the progressing cancer such as pain) offered after lenvatinib or sorafenib.

Cabozantinib works in a similar way to sorafenib and lenvatinib and is proposed as a treatment option following these two therapies for DTC patients where RAI treatment is no longer stopping the disease from progressing (RAI-refractory). The European Society of Medical Oncology (ESMO) updated their guidelines in April 2022, stating that cabozantinib is a potential choice for second-line treatment, but that the optimal sequence cannot be determined currently[17].

The diagram below outlines the treatment overview for RAI-refractory DTC, including cabozantinib and adapted from the ESMO and NICE recommendations[17-18].

Cabozantinib should not be prescribed if the patient is allergic to cabozantinib or any of its ingredients. Cabzantinib can affect the way some other medicines work so that they are less effective or increase the risk of side effects. Also, some medicines can affect the way cabzantinib works and likewise increases the risk that cabozantinib may not work as well or that the risk of side effects is higher. Therefore, any other medicines that a patient may already be taking or about to start taking need to be checked against cabozantinib to see if they could affect each other. This would be done by the doctor or pharamcist usually.



2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

Response: McIntyre et al. (2018) investigated the quality of life in a cohort of UK thyroid cancer patients through a patient-doctor thyroid cancer forum. The forum did not specifically focus on RAI-refractory DTC patients; however it provided an opportunity to highlight areas of improvement for patient care. The forum was founded by the Butterfly Thyroid Cancer Trust and study found that the average quality of life of this group of thyroid cancer patients was lower than the UK population average[19].

Ipsen has not carried out any PBE about patient needs and disease experiences for this HTA submission.

SECTION 3: The treatment

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities.

If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.

Response: Cabozantinib works in a similar way to lenvatinib and sorafenib, by binding to receptor tyrosine kinases that have been upregulated in the cancer cell signalling process. Through this reversible binding, cabozantinib inhibits cellular activity and prevents further growth of the cancer. Cabozantinib is found to be a potent inhibitor at specific receptor tyrosine kinases, all known to be important in the pathology of thyroid cancer, specifically DTC[20].

Useful links to European Medicines Agency documents:
https://www.ema.europa.eu/en/documents/overview/cabometyx-epar-medicine-overview_en.pdf
https://www.ema.europa.eu/en/documents/product-information/cabometyx-epar-product-information_en.pdf

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines?

- Yes / No

If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together.

If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects.

If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.

Response: No, cabozantinib will be used as a monotherapy.

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for.

How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?

Response: Cabozantinib is an oral medicine, with one 60 mg tablet to be taken once daily. Treatment should continue until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs. Management of possible side effects and drug reactions may require treatment breaks or dose reductions of cabozantinib. When dose reductions are necessary, it is recommended to take one 40 mg tablet daily and then 20 mg daily if necessary. Dose reductions are recommended if the side effects are persistent and become serious. The drug side effects of cabozantinib are typical of this class of medicine (TKIs), which also include lenvatinib and

sorafenib and is consistent with the known safety profile of cabozantinib in other disease areas for which it is a licensed treatment[21-22].

3d) Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

Response: The COSMIC-311 global clinical trial tested the efficacy and safety of cabozantinib in adult patients with RAI-refractory advanced DTC, who have progressed during or after prior systemic therapy[23].

Study	**COSMIC - 311**
Study design	Randomised, double-blind, placebo-controlled, phase III
Population	Patients with previously treated advanced RAI-Refractory DTC
Settings and locations where the data were collected	A total of 258 subjects were randomised in 161 unique sites by 174 principal investigators in 25 countries in Asia, North America, Europe and the rest of the world.
Completion Date	December 2022 (Start date: October 2018)
Intervention(s)	Oral cabozantinib 60 mg once daily plus best supportive care (BSC)
Comparator(s)	Oral matched placebo once daily plus BSC
Reported outcomes specified in the decision problem	• Overall survival (OS) • Progression-free survival (PFS) • Time to treatment discontinuation (TTD) • Objective response rate (ORR) • Adverse events (AEs) • Health-related quality of life (EQ5D-5L)
All other reported outcomes	Pharmacokinetics
Key Inclusion Criteria	• Histologically or cytologically confirmed diagnosis of DTC, including the following subtypes: • Papillary thyroid carcinoma (PTC) including histological variants of PTC • Follicular thyroid carcinoma (FTC) including histological variants of FTC • Measurable disease according to RECIST 1.1 on CT/MRI performed within 28 days prior to randomization • Must have been previously treated with or deemed ineligible for treatment with Iodine-131 for DTC • Patients must have received at least one prior Vascular Endothelial Growth Factor Receptor (VEGFR)-targeting TKI therapy of either lenvatinib or sorafenib and must have had radiographic progression during treatment or within 6 months after the most recent dose of the VEGFR inhibitor (up to two prior therapies were allowed including, but not limited to, lenvatinib and sorafenib) • Must have experienced documented radiographic progression per RECIST 1.1 per the Investigator during or following treatment with a VEGFR-targeting TKI prior to starting the next anticancer therapy (which may have been treatment in COSMIC-311) • Age – 16 years and older (Adult, Older Adult) • Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 or 1

Key Exclusion Criteria	• Prior treatment with any of the following:
	  ○ Cabozantinib
	  ○ Selective small-molecule BRAF kinase inhibitor (e.g., vemurafenib, dabrafenib)
	  ○ More than 2 VEGFR-targeting TKI agents (e.g., lenvatinib, sorafenib, sunitinib, pazopanib, axitinib, vandetanib)
	  ○ More than 1 immune checkpoint inhibitor therapy (e.g., programmed cell death-1 [PD-1] or programmed cell death ligand 1 [PD-L1] targeting agent)
	  ○ More than 1 systemic chemotherapy regimen (given as single agent or in combination with another chemotherapy agent)
	• Receipt of any type of small molecule kinase inhibitor within 2 weeks or 5 half-lives of the agent, whichever was longer, before randomisation
	• Receipt of any type of anticancer antibody or systemic chemotherapy within 4 weeks before randomization
	• Receipt of radiation therapy for bone metastasis within 2 weeks or any other radiation therapy within 4 weeks before randomisation
	• Subjects with clinically relevant ongoing from prior radiation therapy that had not completely resolved were not eligible

Clinical Data Cuts	Data from two data cuts are available from the COSMIC-311 trial: 19th August 2020 and 8th February 2021 as clinical cut-off 1 (CCO1) and clinical cut-off 2 (CCO2), respectively.

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

Response: Cabozantinib was found to deliver extended survival and delayed disease progression in patients with advanced RAI-refractory DTC who have received prior therapy, in the COSMIC-311 trial. At the cut-off date for the final analysis, there was high maturity of progression-free survival (PFS) with the trial showing that cabozantinib reduced the risk of disease progression or death by 78% at both CCO1 and CCO2 timepoints. Overall survival (OS) was not a primary endpoint in COSMIC-311 and the trial was not designed to support statistically significant OS. However, the analysis of both CCO1 and CCO2 supported the trend of longer OS for cabozantinib patients in the trial, with the stratified HRs being 0.76 and 0.54 for CCO2 and CCO1, respectively. The descriptive analyses of OS did not show a statistically significant benefit, however did show a trend in benefit for cabozantinib versus placebo in COSMIC-311, despite crossover between the groups. The benefits of cabozantinib were accompanied with a manageable safety profile, similar to that of its drug class.

A feature of the trial design for COSMIC-311 was the permitting of crossover of subjects in the placebo arm to the cabozantinib arm upon disease progression. Data was collected independently when these patients switched to treatment, and therefore it is necessary to mitigate bias in the OS results (see sections B.2.4.2.2 and B.2.6.4 for further information). The ability of patients to switch from placebo to active treatment with cabozantinib on disease progression may also partly

explain why no significant improvement in OS is found from the trial. But it is important to note that COSMIC-311 was not designed to capture improvement in OS therefore it is difficult to interpret OS results.

Additionally, the crossover subjects may have had different prognosis than non-crossover subjects, impacting the health-related quality of life (HRQoL) captured in COSMIC-311. Understandably for patients, HRQoL is an important outcome however cannot be adjusted for in this trial due to the suspected differing of prognosis of these crossover patients. Further discussion on the HRQoL results can be found in section B.2.6.5.

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs).**

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

Response: HRQoL was measured throughout COSMIC-311 using the EQ-5D-5L method, where patients are asked to complete the questionnaire before receiving treatment or control, before collecting the results every 4 weeks until week 25 and then every 8 weeks thereafter. Questionnaires were discontinued after disease progression and for patients who transitioned to the crossover phase. On all dimensions of the EQ-5D-5L, changes from pre-treatment to post-treatment in both the cabozantinib and placebo arms did not show any statistically or clinically meaningful difference meaning that treatment of RAI-refractory DTC with cabozantinib does not show a quality-of-life deterioration for patients despite any side effects they may experience. This is not consistent with other oncology treatments, as usually an early deterioration of QoL is expected because of drug side effects.

The McIntyre study reported PROs about the implications of diagnosis and surgery, including: scarring, fatigue, forgetfulness, weight gain and depression since diagnosis. However, this study is not specifically for RAI-refractory patients, so other PROs might be more relevant to this group of patients. The patients who attended the patient/doctor thyroid cancer conference were not randomly selected to attend and they may not be representative of all UK patients with DTC, however this was not specific to RAI-refractory DTC patients[19].

Patient preference information for DTC shows that patients are focused on palliative care, rather than the effect and benefit of the TKI drug class. Koot et al. (2021) identified the needs, preferences and values of patients with DTC in The Netherlands. This study also considered the TKI treatment group, however there is limited data to compare the DTC-specific findings. It is suggested that due to the metastatic disease, the focus for patients is instead on survival whilst maintaining a good QOL, instead of the focus of clinicians on PFS[24].

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

Response: The main side effects of cabozantinib are similar to other medicines in its drug class, with the COSMIC-311 trial showing that cabozantinib has a manageable safety profile in RAI-refractory DTC. The most frequently reported side effects include diarrhoea, hand-foot syndrome, hypertension and fatigue.

The number of side effects leading to dose reductions was 57%, side effects leading to dose interruption 69% and dose modification 75%. The initial dosing for this indication is 60 mg a day, however the dose can be reduced to 40 mg first, before further decreasing to 20 mg. This enables clinicians to start patients on the most effective dose, however, have the flexibility to reduce the dose after seeing initial benefit but managing toxicity/severe side effects. Clinical experts consulted by Ipsen state that most severe side effects with cabozantinib can be managed by outpatient or remote consultation without the need to be admitted into a hospital bed. The rate of treatment discontinuation due to treatment-related side effects in the COSMIC-311 trial was reasonably low at 8.8%.

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration

Response: RAI-refractory advanced DTC patients in the cabozantinib arm of the COSMIC-311 trial demonstrated a clinically meaningful and statistically significant prolongation of PFS. The results also show that there was a trend of improvement in OS with cabozantinib, with reduction of the thyroid cancer lesion size and disease stabilisation also shown in favour of cabozantinib treatment.

Additionally, cabozantinib is an oral tablet which is only needed to be taken once a day. The dose can be reduced if patients do show severe side effects, which also only need be taken once a day. The COSMIC-311 trial showed that the treatment of cabozantinib in patients did not deteriorate health-related quality of life either through the EQ-5D results.

Currently, there are no options post sorafenib or lenvatinib for advanced RAI-refractory DTC. The only available support is BSC, therefore having the option of a further therapeutic as a second-line option will support the unmet need of these patients.

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response: • Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers? • Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration • What is the impact of any disadvantages highlighted compared with current treatments
Response: Cabozantinib is a TKI with an associated side effect profile, however this is not unmanageable and cabozantinib is an established treatment for other kinds of cancer such as kidney and liver cancer. Clinicians understand how to manage the toxicities associated with the TKI drug class, however this will be considered by clinicians when prescribing cabozantinib to patients.

3i) Value and economic considerations

Introduction for patients: Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model. In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on: • The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by patients; were any improvements that would be important to you missed out, not tested or not proven?) • If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)? • How the condition, taking the new treatment compared with current treatments affects your quality of life.
Response: As part of the NICE submission a cost-effectiveness analysis (CEA) was conducted to show the value for money of the introduction of cabozantinib as a treatment option for DTC patients compared to BSC. As part of a CEA an incremental cost-effectiveness ratio (ICER) is determined. NICE has pre-defined ICER thresholds to determine whether a medicine is cost-effective. The results of the CEA show that cabozantinib is a cost-effective treatment for RAI-refractory DTC as the ICER falls within NICE thresholds. The structure of this CEA is a partitioned survival model over a patient's lifetime. This is a standard structure for an oncology CEA. The model is separated into 3 health states; PFS, progressed-disease (PD), and death. The proportion of patients in each health state at a given time is determined by the results from the endpoints in the COSMIC-311 trial, PFS and OS. The results from the COSMIC-311 trial show that cabozantinib has a significant effect in delaying the progression of the disease and shows a trend in extending the life of patients, relative to placebo. The COSMIC-311 trial has a short follow-up time, with a median* follow up of approximately 10 months. This means that in order to model the length of progression-free

disease and the survival of patients, it is necessary to extrapolate PFS and OS. Extrapolating*** the data from the trial gives an estimate of the proportion of patients who have progression-free disease and who are still alive at future time points.

Another key health outcome measured in a CEA is the quality-adjusted life year (QALY). The QALY is derived from a utility** measurement of the quality of life for patients usually measured directly from the clinical trial or derived from external sources. The utility measurement used to calculate the QALY is from the literature[25]. Fordham et al. (2015) was used for the utility measurement as clinicians[26] felt the utility values were most representable of clinical practice, justifying that a patient's quality of life after progression of disease would be low. In this CEA, there are different utility values based on the 3 health states: 0.87 for PFS, 0.52 for PD, and 0 for death.

The QALY is the key component of a CEA as it informs the ICER. An incremental QALY combines the proportion of extended life for patients relative to the current standard of care and the differences in the quality of life. This incremental QALY informs the ICER which is a ratio between the incremental cost of implementing the new treatment and the incremental QALY.

The additional cost of monitoring and adverse reactions as a result of cabozantinib treatment is considered also. Patients who initiate treatment on cabozantinib are recommended to have an electrocardiogram (ECG) while on treatment (approximately twice a year), these costs are included in the CEA. There is a possibility of serious side effects while on treatment with cabozantinib and these may require medical intervention which would incur a cost. Although, generally side effects can be controlled via virtual appointments with a clinician and thus do not have a significant burden on financial resources for patients or the NHS. The most significant additional cost of cabozantinib is the drug cost itself. Ipsen has in place a patient access scheme (PAS) in order to provide better value for money to the NHS and ensure patients have access to treatment.

As the CEA is based on a clinical trial with limited follow-up, assumptions need to be made to provide a model that is most fitting to clinical practice. A key assumption is made on the OS for the BSC patients in the CEA. Clinical input[26] suggested that the OS extrapolations were overestimating the length of life for the BSC patients and stated that it would not be plausible that after 5 years any patient would be expected to be alive with BSC. The CEA considers this clinical input and does not calculate any costs or QALYs after 5 years.

Overall, the CEA results show that cabozantinib extends the life of patients, increases quality of life by significantly delaying progression of disease and incurs additional cost to the NHS relative to BSC. The incremental cost-effectiveness ratio is £28,148 for cabozantinib compared to BSC.

In 2022 NICE updated its processes to include a severity modifier for ICER thresholds, displacing the previous end of life criteria. A severity modifier gives higher ICER thresholds for the most severe diseases. Given the short length of life expected with no systemic treatment for RAI-refractory DTC after progression of lenvatinib or sorafenib, the disease area will have a severity modifier of 1.2 applied to the ICER threshold for this appraisal.

*Median: relating to a value or quantity lying at the midpoint of a frequency distribution of observed values or quantities, such that there is an equal probability of falling above or below it.

**Utility: Health utility is a measure of the preference or value that an individual or society gives a particular health state, with 1 being perfect health and 0 being death.

***Extrapolation: In health economics, extrapolating OS and PFS is required to give a prediction of future OS and PFS. Extrapolation is the action of estimating or concluding something by assuming that existing trends will continue.

3j) Innovation

NICE considers how innovative a new treatment is when making its recommendations.
If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f)

Response: Cabozantinib represents another option for advanced RAI-refractory DTC patients, introducing a further line of treatment for clinicians to introduce to prevent further tumour progression. Cabozantinib is not a new medicine and is used in kidney cancer and liver cancer, however this enables oncologists to understand the prescribing needs.

3k) Equalities

Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged.
Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics

More information on how NICE deals with equalities issues can be found in the NICE equality scheme
Find more general information about the Equality Act and equalities issues here

Response: Thyroid cancer European age standardised (AS) incidence rates for females and males combined increased by 175% in the UK between 1993-1995 and 2016-2018. The increase was of a similar size in females and males. Females are much more likely to be diagnosed with thyroid cancer making up 72% of thyroid cancer cases in the UK. The AS incidence for thyroid cancer in females is 8.7 and for male it is 3.6 per 100,000, respectively. Therefore, cabozantinib in DTC will reduce the health inequalities for female thyroid cancer patients[2].

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc.
Where possible, please provide open access materials or provide copies that patients can access.

Response:
COSMIC-311 Information:

- Published clinical trial data available at: https://www.thelancet.com/journals/lanonc/article/PIIS1470-2045(21)00332-6/fulltext
    - This only shows results from data cut-off 1, whereas the CCO2 (data cut off 2) results have not yet been published in full.

- Further information about the clinical trial available at: https://clinicaltrials.gov/ct2/show/NCT03690388

Background Information about Thyroid Cancer:

- NHS information: https://www.nhs.uk/conditions/thyroid-cancer/
- Cancer Research UK: https://www.cancerresearchuk.org/about-cancer/thyroid-cancer/stages-types/types
- Thyroid Cancer Statistics in the UK: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/thyroid-cancer
- British Thyroid Foundation: https://www.btf-thyroid.org/thyroid-cancerleaflet

Further information on NICE and the role of patients**:**

- Public Involvement at NICE Public involvement | NICE and the public | NICE Communities | About | NICE
- NICE's guides and templates for patient involvement in HTAs Guides to developing our guidance | Help us develop guidance | Support for voluntary and community sector (VCS) organisations | Public involvement | NICE and the public | NICE Communities | About | NICE
- EUPATI guidance on patient involvement in NICE: https://www.eupati.eu/guidance-patient-involvement/
- EFPIA – Working together with patient groups: https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf
- National Health Council Value Initiative. https://nationalhealthcouncil.org/issue/value/
- INAHTA: http://www.inahta.org/
- European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

4b) Glossary of terms

Response:

Abbreviations:

BSC – Best supportive care
CEA – Cost-effectiveness analysis
CCO1 and CCO2 – clinical cut-off 1 and 2
DTC – Differentiated thyroid carcinoma
ECG - Electrocardiogram
HRQoL – Health related quality of life
ICER – Incremental cost-effectiveness ratio
OS – Overall survival
PAS – Patient access scheme
PFS – Progression-free survival
QALY – Quality-adjusted life year
RECIST – Response Evaluation Criteria in Solid Tumour
TKI – Tyrosine kinase inhibitor
VEGFR – Vascular Endothelial Growth Factor Receptor

Hazard Ratio: Hazard ratios (HRs) are used in clinical trials to measure survival at any point in a group of patients who have been given a specific treatment compared to the control group given placebo. HRs measure how often a particular event happens in one group compared to how often it happens in another group over time.

ICER: An incremental cost effectiveness ratio is calculated by the difference in cost between the new treatment and the standard of care, divided by the difference in health effects (QALYs).

Median: relating to a value or quantity lying at the midpoint of a frequency distribution of observed values or quantities, such that there is an equal probability of falling above or below it.

QALYs: The quality-adjusted life year is a generic measure of disease burden, including both the quality and the quantity of life lived.

RECIST: RECIST is a standard way to measure how well a cancer patient responds to treatment, based on whether the tumour lesion shrink, stay the same or get bigger.

Utility: Health utility is a measure of the preference or value that an individual or society gives a particular health state, with 1 being perfect health and 0 being death.

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

Response:

1. International Agency for Research on Cancer (2021), United Kingdom. Available at: https://gco.iarc.fr/today/data/factsheets/populations/826-united-kingdom-fact-sheets.pdf?msclkid=05ebee3ea3b911eca463c9ff3df72273 Accessed September 2022.
2. CRUK – Thyroid Cancer Incidence Statistics. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/thyroid-cancer/incidence#heading-Zero Accessed September 2022.
3. CRUK – Types of Thyroid Cancer. Available at: https://www.cancerresearchuk.org/about-cancer/thyroid-cancer/stages-types/types Accessed September 2022
4. Dal Maso L, Tavilla A, Pacini F et al. (2017) Survival of 86,690 patients with thyroid cancer: A population-based study in 29 European countries from EUROCARE-5. European Journal of Cancer 1;77:140-152.
5. Office for National Statistics (2019) Cancer survival in England - adults diagnosed. Accessed March 2022.
6. Filetti S, Durante C, Hartl D, et al. Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2019;30(12):1856-83. doi: 10.1093/annonc/mdz400
7. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;26(1):1-133. doi: 10.1089/thy.2015.0020 [published Online First: 2015/10/16]
8. Schmidt A, Iglesias L, Klain M, et al. Radioactive iodine-refractory differentiated thyroid cancer: an uncommon but challenging situation. Arch Endocrinol Metab 2017;61(1):81-89. doi: 10.1590/2359-3997000000245 [published Online First: 2017/02/23]
9. Haddad RI, Nasr C, Bischoff L, et al. NCCN Guidelines Insights: Thyroid Carcinoma, Version 2.2018. J Natl Compr Canc Netw. 2018;16(12):1429-1440. doi:10.6004/jnccn.2018.0089
10. Livrov R., et al. The Treatment of Advanced Thyroid Cancer in the Age of Novel Targeted Therapies. *Therapy in Practice* 2017. 77:733–745.

11. Babu G, Kainickal CT. Update on the systemic management of radioactive iodine refractory differentiated thyroid cancer (Review). Mol Clin Oncol 2021;14(2):35. doi: 10.3892/mco.2020.2197 [published Online First: 2021/01/14]
12. Durante C, Haddy N, Baudin E, et al. Long-term outcome of 444 patients with distant metastases from papillary and follicular thyroid carcinoma: benefits and limits of radioiodine therapy. J Clin Endocrinol Metab 2006;91(8):2892-9. doi: 10.1210/jc.2005-2838 [published Online First: 2006/05/11]
13. Gild ML, Tsang VHM, Clifton-Bligh RJ, et al. Multikinase inhibitors in thyroid cancer: timing of targeted therapy. Nat Rev Endocrinol 2021;17(4):225-34. doi: 10.1038/s41574-020-00465-y [published Online First: 2021/02/20]
14. Cabanillas ME, Hu MI, Durand JB, et al. Challenges associated with tyrosine kinase inhibitor therapy for metastatic thyroid cancer. J Thyroid Res 2011;2011:985780. doi: 10.4061/2011/985780 [published Online First: 2011/10/19]
15. NHS – Thyroid Cancer: Diagnosis. Available at: Thyroid cancer - Diagnosis - NHS (www.nhs.uk) Accessed September 2022
16. Medicines and Healthcare products Regulatory Agency (MHRA). Cabozantinib Summary of Product Characteristics. Available at: https://products.mhra.gov.uk/substance/?substance=CABOZANTINIB Accessed September 2022
17. National Institute for Health and Care Excellence (NICE). Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine. Technology appraisal guidance [TA535] Published: 08 August 2018.Available at: https://www.nice.org.uk/guidance/ta535 Accessed 04/07/22
18. Filetti S, Durante C, Hartl D, et al. ESMO Clinical Practice Guideline update on the use of systemic therapy in advanced thyroid cancer. Annals of Oncology 2022. 33: 674–684.
19. McIntyre C, Jacques T, Palazzo F, Farnell K, Tolley N. Quality of Life in Differentiated Thyroid Cancer. Int J Surg (2018) 50:133–6. doi: 10.1016/j.ijsu.2017.12.014
20. Yakes FM, Chen J, Tan J, et al. Cabozantinib (XL184), a novel MET and VEGFR2 inhibitor, simultaneously suppresses metastasis, angiogenesis, and tumor growth. Mol Cancer Ther 2011;10(12):2298-308. doi: 10.1158/1535-7163.MCT-11-0264 [published Online First: 2011/09/20]
21. Schlumberger M, Tahara M, Wirth LJ, et al. Lenvatinib versus placebo in radioiodine-refractory thyroid cancer. N Engl J Med 2015;372(7):621-30. doi: 10.1056/NEJMoa1406470 [published Online First: 2015/02/12]
22. Brose MS, Nutting CM, Jarzab B, et al. Sorafenib in radioactive iodine-refractory, locally advanced or metastatic differentiated thyroid cancer: a randomised, double-blind, phase 3 trial. Lancet 2014;384(9940):319-28. doi: 10.1016/s0140-6736(14)60421-9 [published Online First: 2014/04/29]
23. Brose MS, Robinson B, Sherman SI, et al. Cabozantinib for radioiodine-refractory differentiated thyroid cancer (COSMIC-311): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Oncology 2021. 22: 1126–1138.
24. Koot A., Netea-Maier R., Ottevanger P., Hermens R., Stalmeier P. Needs, Preferences, and Values during different treatment decisions of patients with differentiated thyroid cancer. J. Pers. Med. 2021, 11(7), 682; https://doi.org/10.3390/jpm11070682
25. Fordham, Beth A et al. "Health state utility valuation in radioactive iodine-refractory differentiated thyroid cancer." Patient preference and adherence vol. 9 1561-72. 3 Nov. 2015, doi:10.2147/PPA.S90425
26. Ipsen advisory board, August 2022.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Clarification questions

September 2022

File name	Version	Contains confidential information	Date
ID4046 cabozantinib clarification response AIC_CIC_26.10.2022_v1.0_FINAL	**1.0**	**Yes**	**26/10/2022**

Section A: Clarification on effectiveness data

Literature Review

A1. Company's submission (CS) Appendix D1.1, pages 6-13. Please comment on the possible risks involved in running a single search strategy across multiple databases, given the considerable variations in indexing and functionality between these sources.

Response

The ProQuest search engine allows to search Embase and Medline simultaneously. The search strategy is developed in such a way that the subject indexing and limits match both the Embase and Medline databases, so that no relevant publications are missed out. The CENTRAL database was searched by means of the advanced search function on the Cochrane Library homepage. In the Cochrane Library, to search the CENTRAL and CSDR databases, a comprehensive list of search terms for second-line or third-line DTC was used to identify relevant literature. The Cochrane search terms for second-line or third-line DTC consisted of words searched in title/abstract and as indexed terms (i.e. MeSH).

A2. CS Appendix D1.1, pages 6-13. Please clarify if the searches of MEDLINE included Medline-In-Process and Epub-Ahead-Of-Print?

Response

Yes, the full Medline database has been searched, including the Medline-in Process/online ahead of print.

A3. CS Appendix D1.1, pages 6-13. The EAG notes that some of the company's search terms have been truncated using the character (*) e.g., "differentia*" (CS Appendix D, Table 2, line 1) but others e.g., "neoplasm" and "tumor" (same table, line 5) have not been truncated in the same way. Please clarify if the ProQuest interface provides automatic lemmatisation (in order to find plural forms of these terms) and if not, explain the implications for the search yield of failing to retrieve these.

Response

Yes, the ProQuest interface provides automatic lemmatisation of terms. Therefore, the truncation character for terms as neoplasm and tumour has not been implemented in the search strategy. This has only been used for search terms with different variations.

A4. CS Appendix D1.1, pages 6-13. The terms for specific interventions (CS Appendix D, Table 2, lines 11-35) include subject headings and free text terms but no field codes. Elsewhere in the strategy the syntax "TI,AB" (or similar) has been used to indicate in which fields terms are to be searched. Where no field code has been indicated, does ProQuest default to searching all fields, or none?

Response

Yes, when no field code is presented in the search string ProQuest by default searches for the term in all fields.

A5. CS, Appendix D.1.2, page 14. The text states that the search was conducted on the 27th September 2021 and a targeted search was updated up to September 2022 using internal Ipsen databases. Please clarify why a full update search was not conducted up to September 2022.

Response

The search was conducted in September 2021 in anticipation of NICE submission in Q1 2022. However, because of scheduling the submission date ended up being in September 2022. The SLR had been done to support NICE and other country HTA submissions so it was impractical to re-do the search just for the UK. Therefore, a decision was made to conduct a targeted search using internal Ipsen databases and resource. This was made on the basis that Ipsen are following this disease space closely and would certainly be aware of all publications for cabozantinib in DTC.

A6. CS, Appendix D.1.2, pages 15-16, Figure 1 and Table 5. Please clarify why the four COSMIC-311 publications identified by the update search (references 42-45 in the CS appendices) are not reported in the included studies listed in Table 5 or the PRISMA flowchart (Figure 1).

Response

This was an oversight in not updating the PRISMA diagram (Figure 1) and Table 5 in the CS appendices. The results of these four studies i.e. Study design (Table 7), PFS

(Table 8), OS (Table 9), Tumour response (Table 10) and General safety summary (Table 13) were incorporated where available in the CS appendices. We have updated the PRISMA diagram below in

Figure 1 and added the missing four trials that should have been included into Table 5 of the CS appendices (see

Table 1).

Figure 1: Updated PRISMA diagram (relevant to Figure 1 in CS appendices)

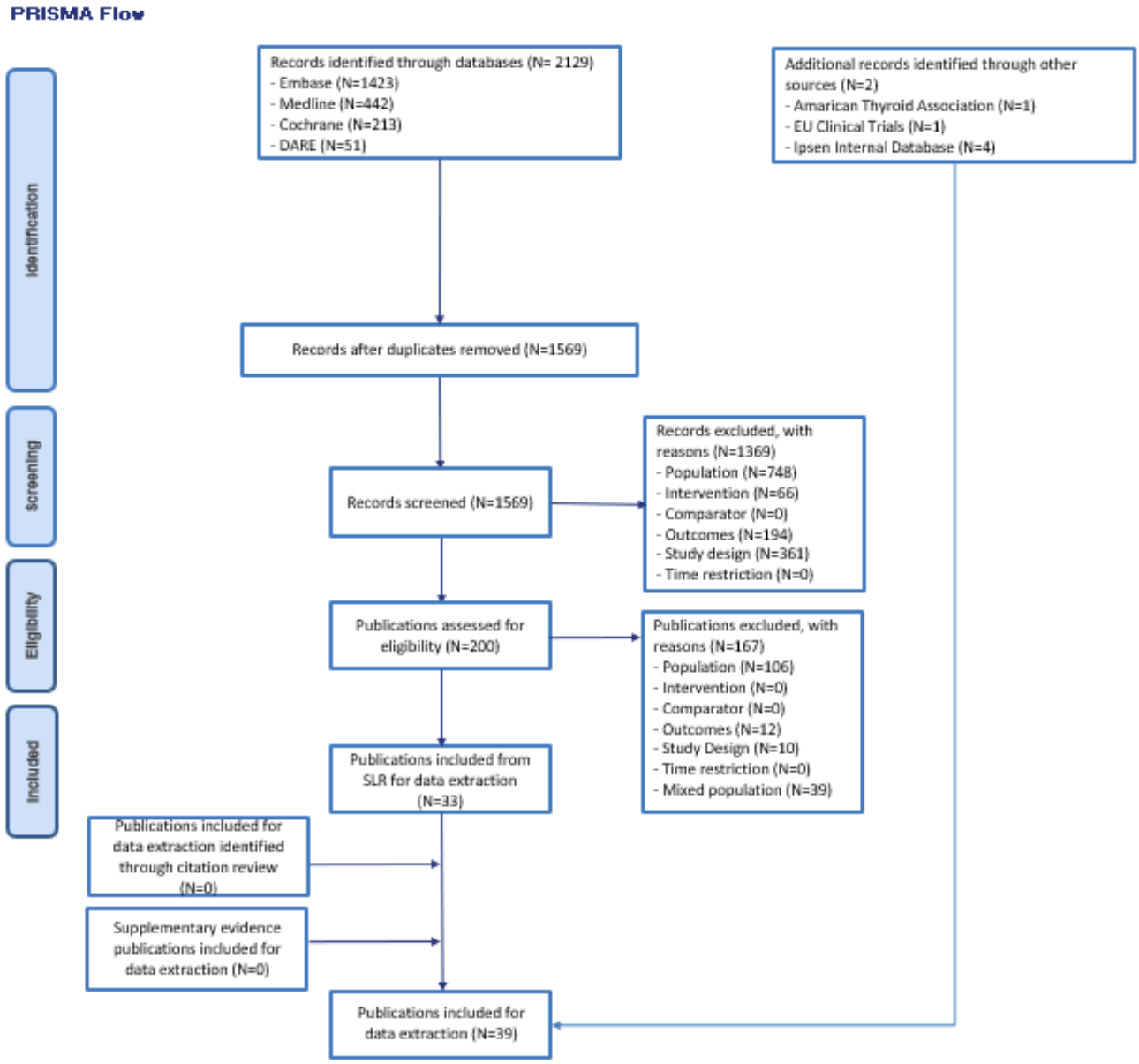


Table 1: Additional included studies (relevant to Table 5 in CS appendices)

Author	Year	Title	Journal	Volume	Pages
Durante et al.[1]	2022	Cabozantinib versus placebo in patients with radioiodine-refractory differentiated thyroid cancer (DTC) who have progressed after prior VEGFR-Targeted therapy: updated results from the phase 3 COSMIC-311 trial and prespecified subgroup analyses based on prior therapy	Endocrine Abstracts	81	Abs OC3.2

Hernando J, et al.[2]	2022	Cabozantinib (C) versus placebo (P) in patients (pts) with radioiodine-refractory (RAIR) differentiated thyroid cancer (DTC) who have progressed after prior VEGFR-targeted therapy: Outcomes in prespecified subgroups based on prior VEGFR-targeted therapy.	Journal of Clinical Oncology	40, no 16 suppl	Abs 6083
Capdevila J, et al.[3]	2022	Cabozantinib versus placebo in patients (pts) with radioiodine-refractory (RAIR) differentiated thyroid cancer (DTC) who progressed after prior VEGFR-targeted therapy: Outcomes in prespecified subgroups based on histology subtypes..	Journal of Clinical Oncology	40, no 16 suppl	Abs 6081
Durante C, et al.[4]	2022	Effect of age on efficacy and safety of cabozantinib vs placebo in patients with radioiodine refractory (RAI-R)-differentiated thyroid cancer (DTC) with progression after VEGFR-targeted therapy: subgroup analysis from Phase 3 COSMIC 311 study.	Endocrine Abstracts	81	Abs OC3.4

A7. CS, Section B.2.2 and CS Appendix D.1. Page 18 of the CS states that *'A SLR did not identify any additional studies relevant to cabozantinib in RAI-refractory advanced DTC'*. Please clarify:

- this statement, given that other relevant cabozantinib trials in RAI-refractory advanced DTC were identified and are listed in Section B.2.8 (e.g., Cabanillas *et al* 2017);
- the disparity between the eligibility criteria for the SLR stated in Appendix D.1.1, Table 1 (any second- or later-line treatment for this population) and the decision problem addressed in the CS (i.e., second-line cabozantinib only);
- why the SLR eligible and included studies listed in Appendix D.1.2 (n=26) are not included in the main clinical effectiveness section, with the exception of COSMIC-311 and NCT01811212.
- why identified studies were excluded from synthesis for a variety of reasons not initially reported in the eligibility criteria (e.g., CS, Section B.2.8 pages 67-68 state that Study NCT01811212 was excluded for having the wrong dose).

Response

- The statement *'A SLR did not identify any additional studies relevant to cabozantinib in RAI-refractory advanced DTC'* is meant to refer to the lack of studies other than COSMIC-311 that fulfil the licensed indication for cabozantinib in DTC and studies also in line with the final scope of this NICE appraisal. The cabozantinib trial in RAI-refractory advanced DTC reported by Cabanillas et al 2017 (NCT01811212) did not use the licensed dosing regimen for cabozantinib. See also the fourth bullet point below and the response to question A9.

- The EAG point regarding the disparity between the eligibility criteria for the SLR stated in Appendix D.1.1, Table 1 (any second- or later-line treatment for this population) and the decision problem addressed in the CS (i.e., second-line cabozantinib only) is noted. On reflection this should have been made clearer. The decision problem we have addressed in this submission is within the licensed indication of cabozantinib and in line with the pivotal COSMIC-311 trial i.e. patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed during or after prior systemic therapy – this could be second or later line treatment. Thus, the SLR aligns with this. The only second-line treatment that has recently been recommended by NICE which could be used to treat RAI-refractory DTC is selpercatinib (TA742). It is recommended for use within the CDF, as an option for treating advanced rearranged during transfection (RET) fusion-positive thyroid cancer in adults who need systemic therapy after sorafenib or lenvatinib. As lenvatinib or sorafenib can only be used first-line in RAI refractory or ineligible patients, and selpercatinib is recommended only within the CDF, the only relevant comparator for cabozantinib is BSC. Because of the treatment algorithm that currently exists based on existing NICE guidance this means cabozantinib is a second-line therapy but it could be used third line as per licensed indication. We are seeking a NICE recommendation based on the licensed indication based on the whole study population from the COSMIC-311 trial. We do state in Table 1 of our CS that the population addressed in our submission is 'adults with locally advanced or metastatic differentiated thyroid carcinoma, whose

disease is refractory to, or who are unsuitable for radioactive iodine, and whose disease has progressed during or after prior systemic therapy.'

- The SLR eligible and included studies listed in Appendix D.1.2 (n=26) are not included in the main clinical effectiveness section, with the exception of COSMIC-311 and NCT01811212 as they are not relevant to the decision problem and scope of comparators in this NICE appraisal, but they could be in other countries. The SLR was conducted to try and satisfy the needs of multiple HTA countries and thus allowed the inclusion of other treatments such as lenvatinib and sorafenib to be included.
- Trial NCT01811212 (Cabanillas et al 2017) was initially included but subsequently excluded when it was realised that the trial allowed the use of a dose (80mg) that is not within the licensed indication of cabozantinib in DTC. The SLR was conducted to try and satisfy the needs of multiple HTA countries. The final SLR report once received by Ipsen UK was reviewed and it was decided at that time that study NCT01811212 did not meet the NICE final scope as it did use the recommended licensed dose of cabozantinib for DTC – see also response to question A9.

A8. CS, Section B.2.8, page 67. Please clarify how the Phase II cabozantinib trial (NCT02041260, reference 58 in CS Section B.2.8 and reference 90 in CS Appendix F, page 104) was identified, as this study does not appear in Tables 5 or 6 of the included or excluded studies from the search in CS Appendix D.

Response

This study was identified during the targeted search of internal Ipsen databases in September 2022 and is listed in the CSR as part of the investigational programme for cabozantinib in DTC. It was decided to reference this in the Appendix F (Adverse Reactions) as this added to the totality of data for cabozantinib safety in DTC but as stated in Appendix F the patient population in terms of efficacy is not relevant to the final scope. It is also mentioned in the main CS for completeness.

A9. CS, Section B.2.8, page 67. Please clarify why trial NCT01811212 was excluded given that only 16% (4/25) of included patients received the escalated 80mg dose of cabozantinib (Cabanillas *et al* 2017). Please also clarify at what point in the review

process the use of a different dose regimen from the licensed dose of 60mg became an exclusion criterion.

Response

As stated above the SLR was conducted to try and satisfy the needs of multiple HTA countries. The final SLR report once received by Ipsen UK was reviewed and it was decided that study NCT01811212 did not meet the NICE final scope nor did use the recommended licensed dose of cabozantinib for DTC.

However, we have reconsidered the point made by the EAG and present the details and results for this study below. As this was an investigator led study and not a company sponsored study we are currently limited in obtaining any further detail of the study beyond that presented in the publication. Because of the potential that this study could support the longer term efficacy of cabozantinib in the health economic model extrapolations for the pivotal phase 3 COSMIC-311 trial [redacted].

Patients with advanced thyroid cancer (papillary, follicular, Hurthle cell, or poorly differentiated) aged ≥ 18 years were eligible if they met the following criteria: measurable disease by Response Evaluation Criteria in Solid Tumor (RECIST) v1.1; RECIST v1.1 progression on prior VEGFR targeted therapy (up to two lines of prior VEGFR-targeted therapy were allowed); RAI-refractory disease as defined by one or more of the following criteria:

1. One or more measurable lesions that did not demonstrate RAI uptake
2. One or more measurable lesions progressive by RECIST v1.1 within 12 months of prior RAI therapy
3. One or more measurable lesions present after a cumulative RAI dose of . 600 mCi

4. Adequate organ function and Eastern Cooperative Oncology Group performance status of 0 or 1.

The multicentre, single arm trial was an investigator led study coordinated by The Academic and Community Cancer Research United (ACCRU) and funded by the NCI (ClinicalTrials.gov identifier: NCT01811212) and the International Thyroid Oncology Group (ITOG).

Cabozantinib was administered orally at a starting dose of 60 mg daily in 28-day cycles. Patients who tolerated cabozantinib with no ≥grade 2 treatment-related AEs could have their dose increased to 80 mg daily. Those patients experiencing ≥grade 2 treatment related AEs had their dose reduced to 40 mg daily (and again to 20 mg daily, if necessary). Treatment was continued until disease progression, unacceptable toxicity, or withdrawal of consent. Of note, patients who had continued clinical benefit (such as symptomatic or tumour marker improvement or decreased tumour burden compared with baseline or limited progression in a nontarget lesion treated with radiation or surgery) were allowed to receive therapy even if they met criteria for progressive disease per RECIST v1.1.

Between September 2013 and January 2015, 25 patients were enrolled by International Thyroid Oncology Group (ITOG) investigators at six centres in United States. Baseline patient characteristics are listed in Table 2. All patients with RAI-refractory DTC had measurable disease and disease progression while receiving at least one line of prior VEGFR-targeted therapy. The majority of patients had aggressive histology (28% PDTC, 20% HTC, 16% FTC), and a high frequency of bone (84%), liver (36%), and brain (20%) metastases was observed. Patients had high tumour burden at study entry, and in addition to RAI, patients were heavily pre-treated with systemic cytotoxic chemotherapy and/or targeted therapies: seven patients had received at least two lines, three had received three lines, and one had received four lines of systemic therapy, including VEGR-targeted therapy. Of five patients with brain metastasis at study entry, all had stable brain metastasis and had discontinued corticosteroids for least for 2 weeks before study entry, four patients had undergone radiation (stereotactic, n = 3; whole brain, n = 1) and one patient had undergone surgery for brain metastasis.

Table 2: Baseline Patient Characteristics

Characteristic	No. (%)
Total No. of patients	25
Accrual period, dates	9/9/2013-1/13/2015
Sex	
Female	9 (36)
Male	16 (64)
Age, years	
Median (range)	64 (41-81)
Race	
White	21 (84)
Black	1 (4)
Asian	3 (12)
Histologic type of thyroid cancer	
Papillary	9 (36)
Poorly differentiated	7 (28)
Hürthle cell	5 (20)
Follicular	4 (16)
Disease stage	
Locally advanced only	1 (4)
Distant metastatic only	14 (56)
Locally advanced and distant metastatic	10 (40)
Metastatic disease sites	
Bone	21 (84)
Brain	5 (20)
Liver	9 (36)
Lung	21 (84)
Lymph nodes	24 (96)
Prior therapies	
Surgery	25 (100)
External beam radiation	19 (76)
Radioactive iodine therapy	21 (84)
Cytotoxic chemotherapy	4 (16)
Targeted therapy (besides VEGFR-TKI)	5 (20)
Only one prior VEGFR-targeted TKI	21 (84)
Sorafenib	12 (48)
Pazopanib	5 (20)
Cediranib	4 (16)
Two prior VEGFR-targeted TKIs	4 (16)
Sorafenib and pazopanib	1 (4)
Sorafenib and cediranib	1 (4)
Axitinib and cediranib	1 (4)
Lenvatinib and pazopanib	1 (4)
Median serum thyroglobulin at study entry, ng/mL, (range)*	1,025 (10-16,000)
Median sum of target lesions at study entry, cm (range)	7.9 (1.5-18.6)

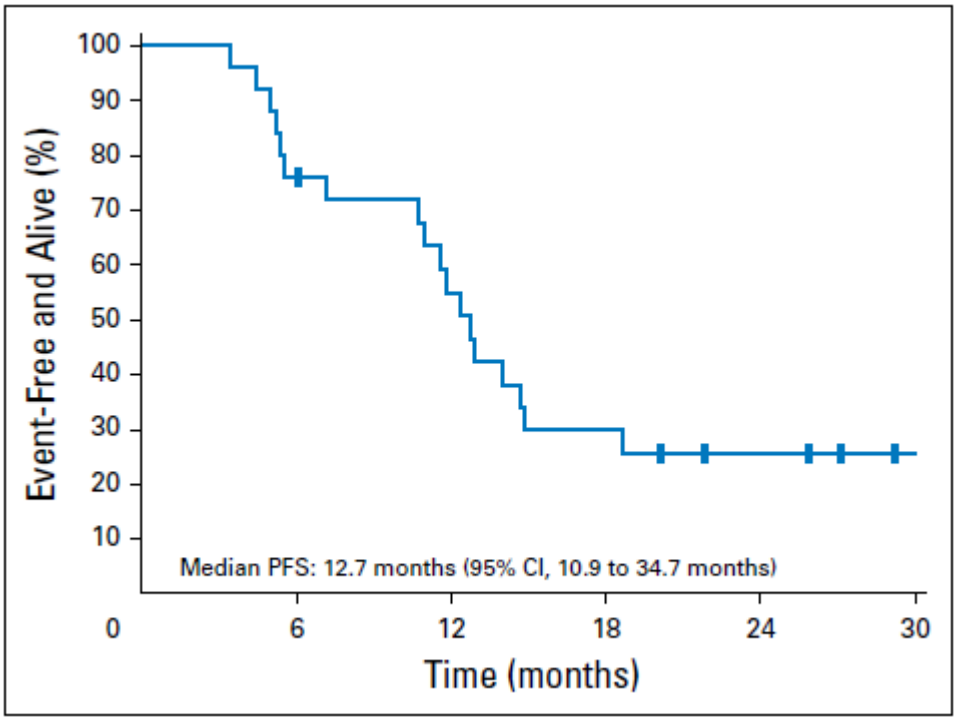
Abbreviations: TKI, tyrosine kinase inhibitor; VEGFR, vascular endothelial growth factor receptor.
*Data were not available for two patients.

Seven patients (28%) were treated at 60 mg/day of cabozantinib, whereas four (16%) had a dose escalation to 80 mg/day and 14 (56%) had a dose reduction to 40 mg/day (n = 6; 24%) or to 20 mg/day (n = 8; 32%).

Median duration of follow-up was 22.8 (95% CI, 21.2 to 30.2) months. The pattern of progressive disease after achieving nadir SD or PR was interesting in that only four patients had a 20% increase in the sum of target lesions compared with the nadir response, whereas target lesions in seven patients maintained reduction or stability. In these seven patients, progression was defined based on new lesions (n = 4) or unequivocal progression of nontarget lesions (n = 3).

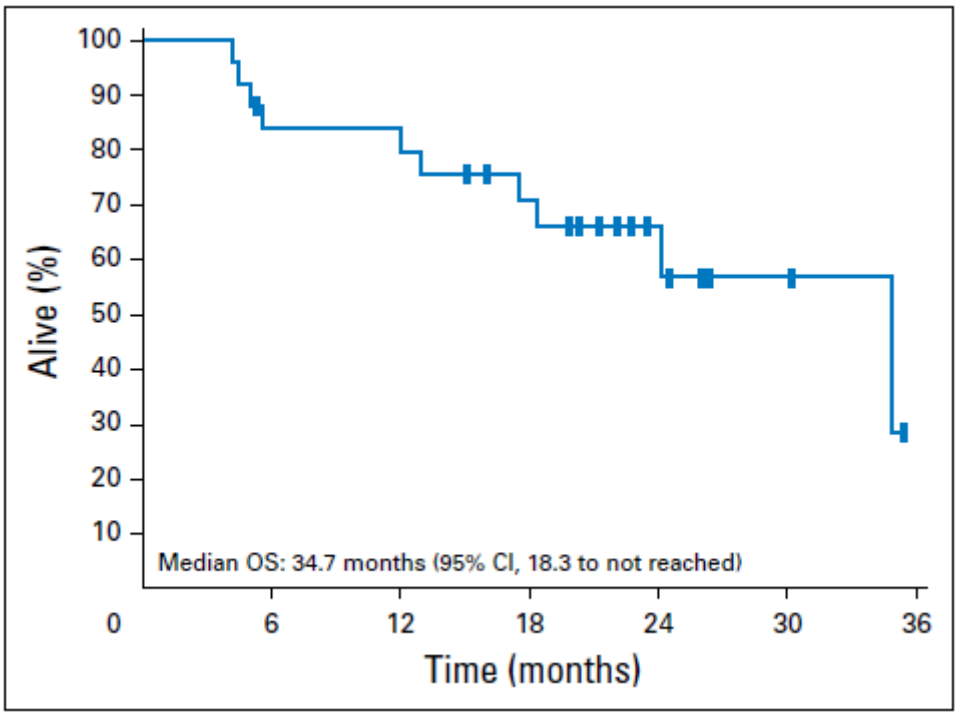
Median PFS was 12.7 (95% CI, 10.9 to 34.7) months (Figure 2), and the estimated PFS at 12 months was 55% (95% CI, 38% to 79%) and 25% (95% CI, 13% to 50%) at 24 months.

Figure 2: Kaplan-Meier curve of progression-free survival (PFS)



Median OS was 34.7 (95% CI, 18.3 to not reached) months (Figure 3), and estimated OS at 12 months was 80% (95%CI, 65% to 97%) and at 24months was 66% (95%CI, 49% to 88%).

Figure 3: Kaplan-Meier curve of overall survival (OS)



A10. CS, Appendix D.1.1, Data extraction and quality assessment sub-section, page 14. Please clarify how many reviewers performed the risk of bias assessments and how disagreements between the reviewers were resolved.

Response

After all relevant publications were identified and received, the relevant data were extracted from the articles. One researcher extracted the data and the second researcher independently reviewed all data extracted for each endpoint. The second reviewer checked the data extraction file for accuracy and completeness, by checking if all data presented in the Excel file corresponded directly with what was presented in the selected articles. Thus, the second reviewer did not only check a data sample but checked all articles. Any discrepancies were resolved by a third reviewer. According to the NICE requirements, as part of any SLR, RCTs and non-randomised/observational studies should be subjected to a Quality Assessment (QA) using a recommended checklist. The quality assessment checklists from the CRD Guidance for Undertaking Reviews in Health Care (2009) was applied for quality assessment.[5] One reviewer conducted the QA of included articles; a second reviewer

checked the accuracy of QA performed for all relevant articles. Any discrepancies were resolved by a third reviewer.

QA was performed for all publications except for conference proceedings, as there would be insufficient methodological data to assess the study quality.

A11. CS, Appendix D.1.2, Summary of findings subsection, pages 44-46. Please clarify the disparity between the statement: *'This SLR identified three relevant RCTs, which were described in nine publications'*, and the 13 publications listed in Table 7.

Response

This is an error. The sentence should have been updated to account for the four publications that were identified from the targeted review. These four publications are listed in the response to question A6. Thus the sentence should be *'This SLR identified three relevant RCTs, which were described in 13 publications'.*

A12. CS, Appendix D.1.2, Summary of findings subsection, page 16, Table 5. Please clarify how the following trial publications were identified, as they are not listed in the included studies reported in Table 5:

- COSMIC-311 publications with references 40 and 41 (not among the four studies identified from the update search either);
- SELECT trial publication reference 46;
- EORTC trial publication reference 47.

Response

There appears to have been some duplication of references in the reference list by the software referencing tool used.

- References 40 and 41 are duplicated in the reference list as reference 11 and 8 respectively – these are captured in Table 5.
- Reference 46 (SELECT trial) is captured as reference 12 in Table 5.
- Reference 47 is captured as reference 39 in Table 5.

Positioning and comparators

A13. PRIORITY. CS, Section B.1.1.3, page 9. The CS states that *"the only relevant comparator for cabozantinib is best supportive care (BSC)."* However, the minutes of the 2022 Ipsen clinical advisory board meeting state that [redacted] Please explain why a clinical and economic comparison has not been made between cabozantinib and continued lenvatinib given after patients have progressed on this treatment. Please explore whether sufficient evidence exists to inform such a comparison.

Response

In our submission and as outlined in the final scope, best supportive care (BSC) is the only relevant comparator. NICE guidance (TA535) for radioactive iodine-refractory (RAI-R) DTC recommend sorafenib and lenvatinib only when:[6]

- they have not had a tyrosine kinase inhibitor before or
- they have had to stop taking a tyrosine kinase inhibitor within 3 months of starting it because of toxicity (specifically, toxicity that cannot be managed by dose delay or dose modification).

NICE recommendations have restricted access to sorafenib and lenvatinib in scenarios where they have not received a prior tyrosine kinase inhibitor (TKI). This means that following progression on a TKI there are no treatment options for patients (except selpercatinib for a subset of patients with evidence of RET mutation). In an advisory board, clinicians explained that it is likely patients will be kept on 1L treatment after progression as the patient may get clinical benefit.[7]

This submission has not included lenvatinib or sorafenib as a comparator as there is no clinical evidence of health outcomes of continued lenvatinib or sorafenib after progression. The SELECT trial (lenvatinib) included patients in 2L, post treatment with a TKI, and patients who had received no previous TKI.[8] The study protocol stated that patients discontinue treatment upon radiographic progression. To include lenvatinib

as a comparator in the context of the treatment landscape is not feasible as there is no clinical evidence to support it.

Cabozantinib is likely to offset costs due to patients continuing on lenvatinib treatment, therefore the ICER is likely overestimated in the submission. Nevertheless, it is too uncertain to include lenvatinib as a comparator due to the limited clinical evidence.

A14. CS, Section B.3.2.1, page 86. One of the inclusion criteria for the economic model population includes *"Previously treated with at least one of the following VEGFR-targeting TKI agents for DTC: lenvatinib or sorafenib."* Please clarify if a positive recommendation is being sought in people who have had exactly one prior VEGFR-targeting therapy, or at least one prior VEGFR-targeting therapy.

Response

We are seeking a positive recommendation in line with the licensed indication of cabozantinib i.e. patients with locally advanced or metastatic differentiated thyroid carcinoma (DTC), refractory or not eligible to radioactive iodine (RAI) who have progressed during or after prior systemic therapy – this could be second or later line treatment and include prior VEGFR-targeting therapy.

See also part of the response to question A7.

A15. PRIORITY. CS, Section B.3.3.4, page 105. The CS states *"Upon applying TTD in the model a rule was also applied whereby the TTD curve could not exceed the PFS curve for cabozantinib since patients are assumed to discontinue treatment upon progression as per the SmPC."* However, Section 4.2 of the SmPC for cabozantinib states that *"Treatment should continue until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs"* - this suggests that continued post-progression treatment is permitted under the licence. Please clarify if the company is seeking a positive NICE recommendation for the use of cabozantinib only up to the point of progression.

Response

Ipsen is seeking a recommendation as per the SmPC - "*Treatment should continue until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs".*

In the cost effectiveness model, the TTD curve could not exceed the PFS curve to reflect COSMIC-311 trial. Only [redacted] of patients in the cabozantinib arm transitioned to open label cabozantinib post progression. The TTD analysis from the trial is uncertain given the high censoring and low number of patients at risk after approximately 8 months. Few patients drive the separation between the TTD and PFS KM curve.

The TTD analysis had missing data. Two different assumptions were made to incorporate the missing data into the KM data: TTD was linked to the date of progression (green line) or the date of the last known dose (blue line) – see Figure 4. This figure shows that TTD based on last known dose is under the PFS curve.

Usually, TTD is based on the date of last dose but in absence of survival analyses performed on this data at this time we have taken a conservative approach in anchoring TTD to PFS by modelling TTD based on the date of progression (green line) and capping this TTD data by PFS.

Figure 4: [redacted]

Clinical Effectiveness

A16. CS, Section B.2.3.3, page 26. Please clarify (and substantiate with reference to the published literature) the statement: *"The demographic and baseline characteristics*

[of COSMIC-311] are representative of DTC epidemiology, with the median age at 65 years for the cabozantinib arm and 66 years for placebo."

Response

In the SELECT pivotal trial for lenvatinib in RAI-R thyroid cancer, the baseline characteristics of the ITT population state the median age was 64 in the lenvatinib group and 61 in the placebo group.[8] In the DECISION pivotal trial for sorafenib in locally advanced or metastatic RAI-R DTC, the median age in the sorafenib arm was 63 years old and 63 years old in the placebo arm of the ITT population.[9] Therefore, we believe that the median age of 65 and 66 for the cabozantinib and placebo arms, respectively, are representative of the DTC epidemiology and in line with other pivotal trials in RAI-R DTC. Additionally, Cancer Research UK statistics state that the incidence rates for thyroid cancer in the UK are highest in people aged 65 to 69 (2016-2018), with the age-specific incidence rates in females rising sharply from around age 10-14, reaching a peak at ages 45-49, then declining steadily before dropping sharply from age 65-69.[10]

Ipsen has a compassionate use program in the UK so patients eligible for cabozantinib in RAI-R DTC can access treatment. The median age of the patients from the requests received ([redacted]) is [redacted] years and the mean age is [redacted] years.

A17. CS, Section B.2.3.3, Table 6, page 28. The table indicates that 1 patient in the cabozantinib group of the COSMIC-311 trial had not previously received any VEGFR-TKI therapy. Why was this patient eligible for recruitment into the trial?

Response

[redacted]

A18. CS, Section B.2.5, Table 10, page 39. Please clarify if patients in the two arms of COSMIC-311 were similar for the prognostic factor of tumour burden/volume.

Response

We do not have the data to answer this question.

A19. CS, Section B.2.10.3, page 71. The text states *"Grade 3/4 adverse events had a low incidence at approximately 5%."* However, Table 17 reports overall frequencies of Grade 3/4 AEs for cabozantinib and placebo at CCO2 of 62% and 28%, respectively. Please clarify if the quoted text is incorrect.

Response

Apologies this should be made clearer. This statement is referring to individual grade 3 or 4 adverse events in Table 18 of the CS were overall low in incidence but recognise this is quite a general statement.

A20. CS, Section B.2.3.1.1, page 23. Are subsequent data-cuts of COSMIC-311 expected beyond CCO2? If so, please provide details of when these data-cuts are expected to become available.

Response

There are no further COSMIC-311 data-cuts planned beyond CCO2.

Section B: Clarification on cost-effectiveness data

Survival analysis and treatment switching analysis

B1. CS, Section B.3.3.3, pages 96-104. The CS describes the use of six standard parametric survival models for PFS and OS. Please explore whether flexible parametric models (such as restricted cubic spline models) could provide more clinically plausible predictions of OS for the cabozantinib and BSC groups.

Response

The use of flexible parametric models, despite their increased use for oncology models, would not offer an improvement to the current limitations in the model, and are therefore not deemed appropriate. As the Kaplan-Meier (KM) curves from the COSMIC-311 trial are relatively short, the extrapolation for these curves is subsequently long. Even with flexible survival modelling, as the curves are modelled independently, we still anticipate the OS curve tails for cabozantinib and BSC to cross, producing unrealistic predictions. This is due to the flexible parametric models anticipated to model more closely to the observed data for which there is a sudden

drop in the cabozantinib observed KM data and a flattening of the BSC KM, due to low patient numbers, expected to produce crossing curves which have been validated as being unrealistic by clinicians at an advisory board held in August 2022.[7]

Additionally, studies have shown that median OS of patients who did not receive salvage therapy for radioiodine-refractory (RAIR) DTC after progressing from a single agent TKI ranged between 10 months and 22 months[11,12] and that the consensus of expert opinion from the Ipsen advisory board[7] from the cost-effective modelling survival analysis that no patients would be expected to be alive at 5 years. Therefore using flexible modelling is still anticipated to result in unrealistic curves modelling survival beyond 5 years.

The Company have updated the economic model to improve the curves which better reflect clinical opinion. This has been performed by incorporating CCO1 data for OS for the BSC arm for the base case. More details on this update can be found in response to B6. Observed data in the BSC arm from CCO1 is in line with expert predictions for OS and improve the extrapolated curves such that the KM curves no longer cross.

B2. PRIORITY. CS, Section B.3.3.3, pages 96-104. Please provide plots showing the empirical/unsmoothed and smoothed hazard functions for the data used in the analysis for PFS, OS and TTD. Please also plot the modelled hazards of each of the parametric survival models for PFS and (RPSFTM-adjusted) OS on top of the empirical and smoothed hazard.

Response

The requested plots are provided below.

The smoothed hazards for cabozantinib OS and PFS follow the observed hazards well for most of the observed period (Figure 5 for OS and Figure 9 for PFS). However, none of the parametric models are able to fit well to the steep upward shape of the hazards towards the end of the observed period (Figure 6 for OS and Figure 10 for PFS).

For placebo OS (RPFST-adjusted), the smoothed hazard exhibits a slight negative slope over time. This is not necessarily conveyed by the unsmoothed hazard, which

looks relatively flat (Figure 7). Some of the parametric models are able to fit the shape of the smoothed hazard fairly well (Figure 8).

For placebo PFS, the smoothed and unsmoothed hazards differ in their level and shape (Figure 11). Several of the parametric models are able to fit the shape of the smoothed and unsmoothed hazards reasonably well (Figure 12).

The smoothed hazards for cabozantinib TTD follow the observed hazards well for most of the observed period (Figure 13). The unsmoothed hazard is relatively flat. While the parametric models are not able to perfectly fit the shape of the smoothed hazard, some of them have a rather flat shape, as the unsmoothed hazard (Figure 14).

Figure 5. Smoothed and unsmoothed hazards for Cabozantinib OS

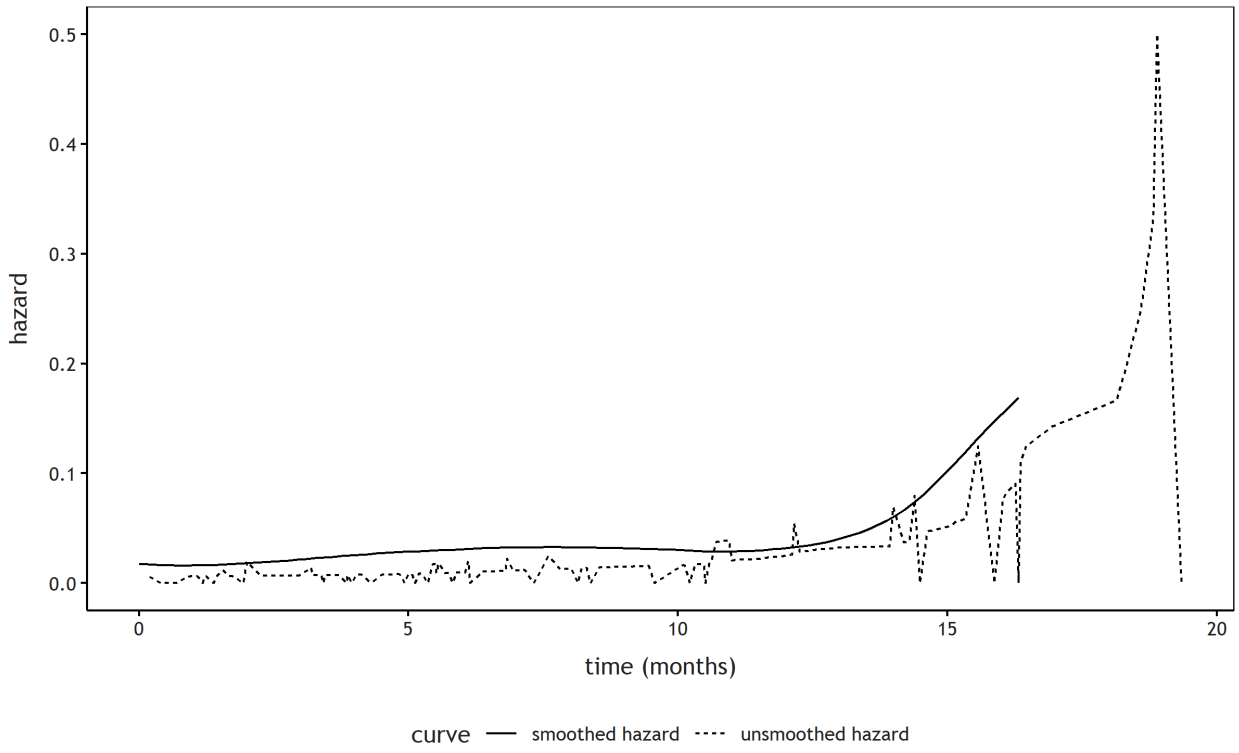


Figure 6. Modelled hazards and smoothed hazards for Cabozantinib OS

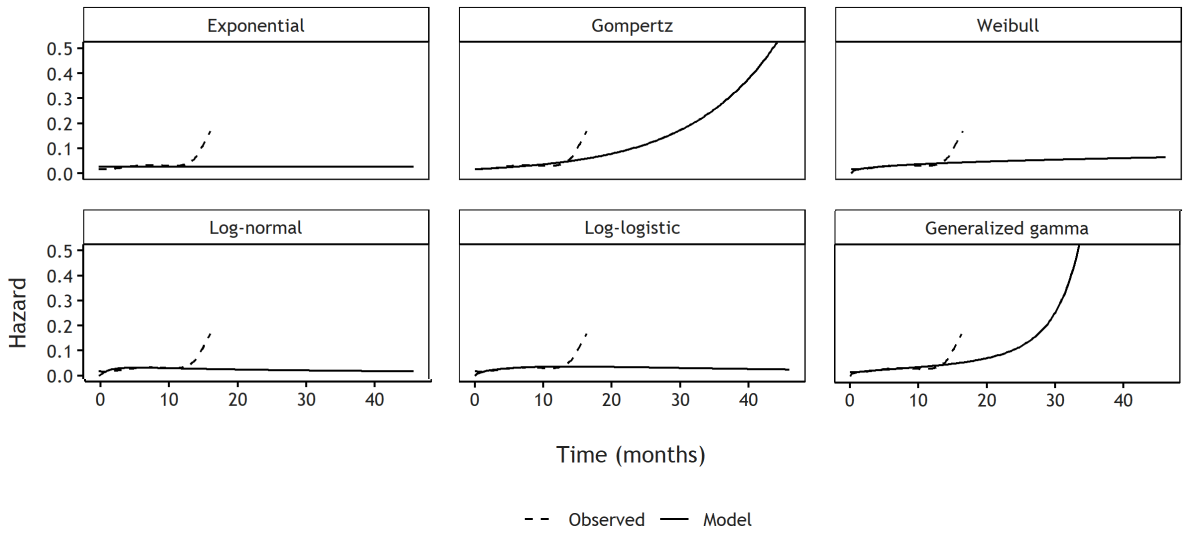


Figure 7. Smoothed and unsmoothed hazards for Placebo OS (RPFST adjusted)

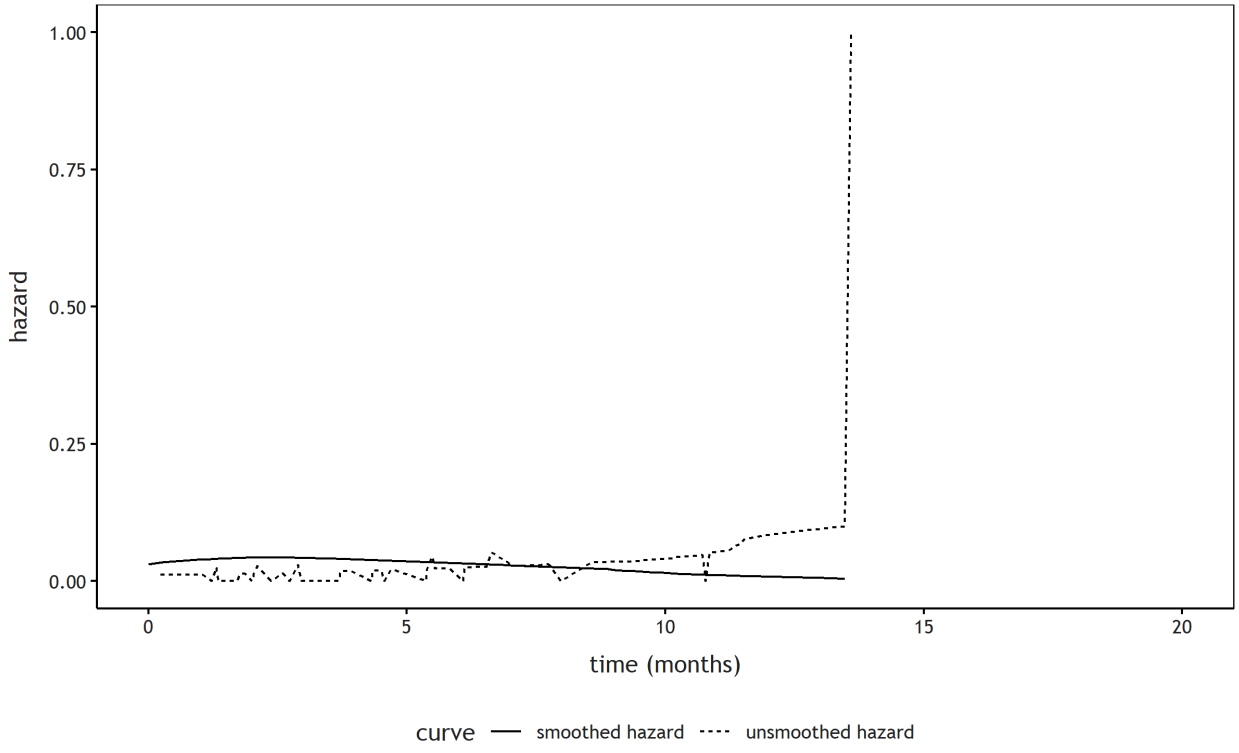


Figure 8. Modelled hazards and smoothed hazards for Placebo OS (RPFST adjusted)

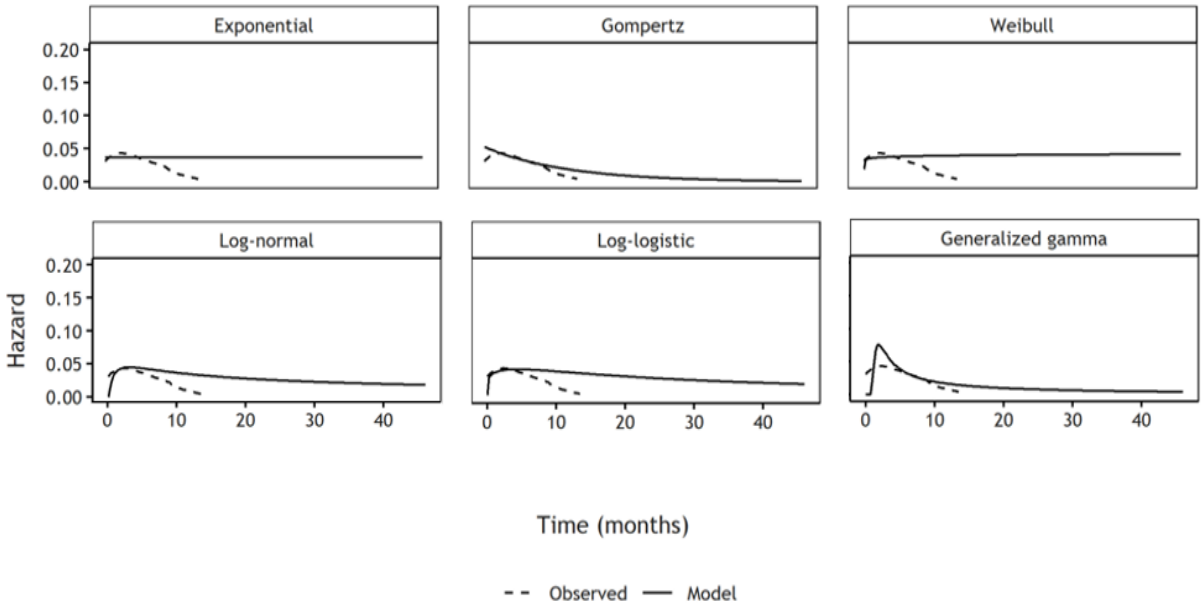


Figure 9. Smoothed and unsmoothed hazards for Cabozantinib PFS

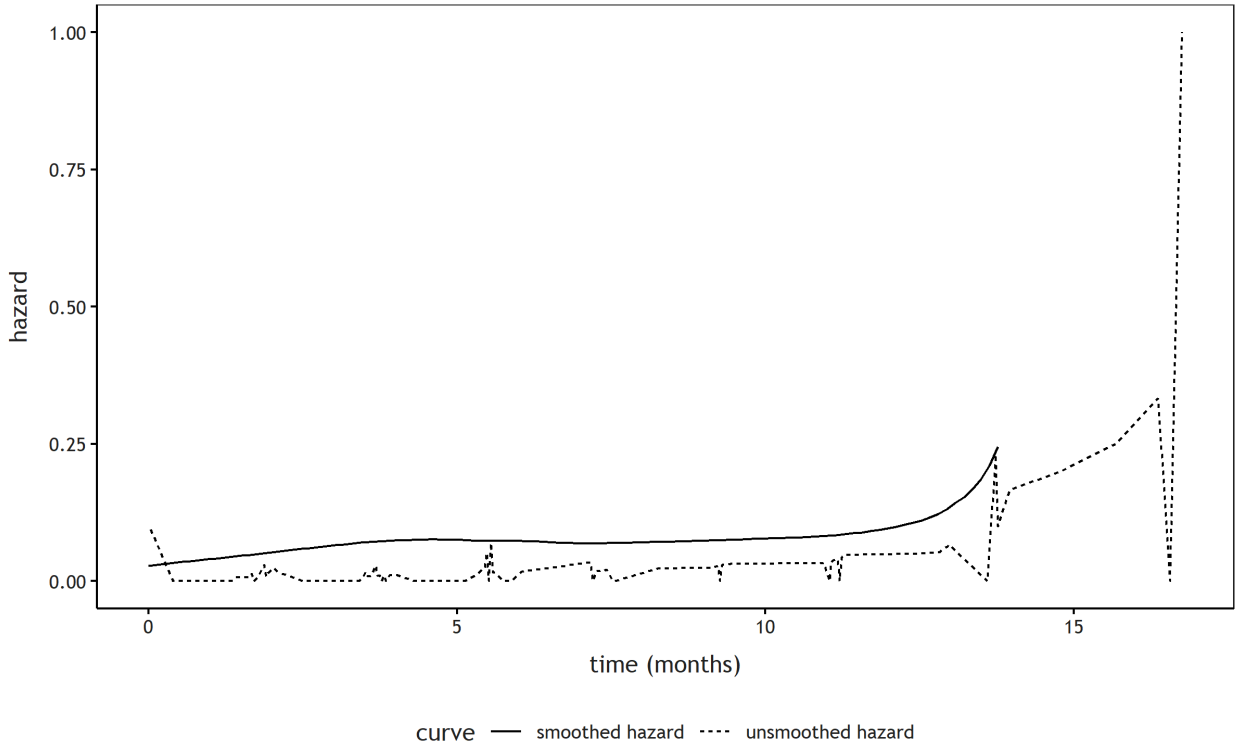


Figure 10. Modelled hazards and smoothed hazards for Cabozantinib PFS

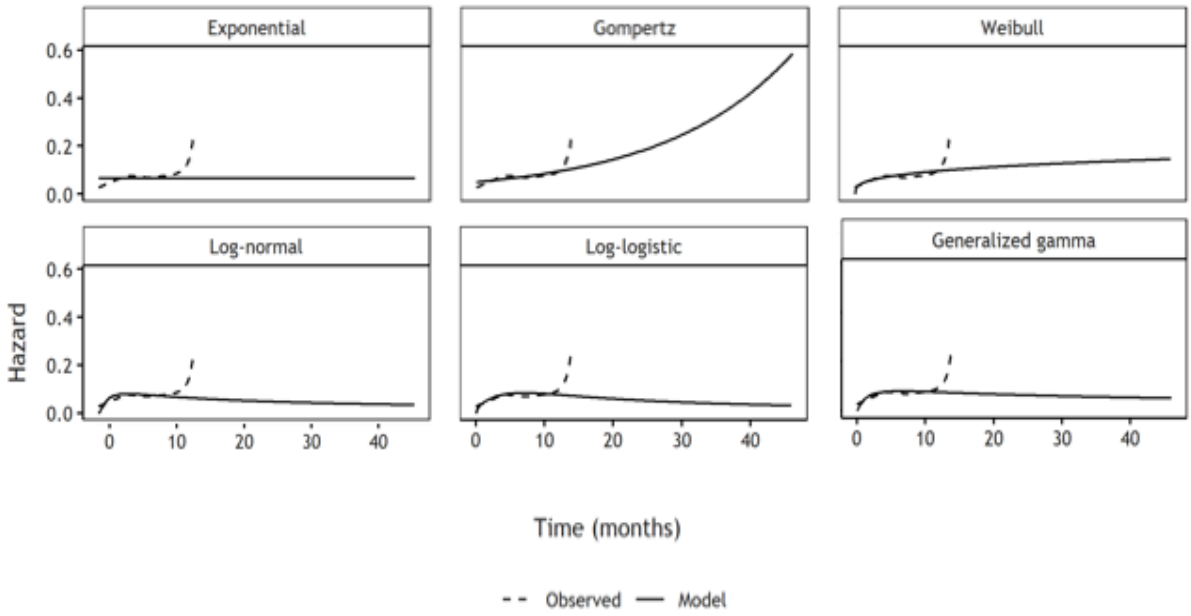


Figure 11. Smoothed and unsmoothed hazards for Placebo PFS

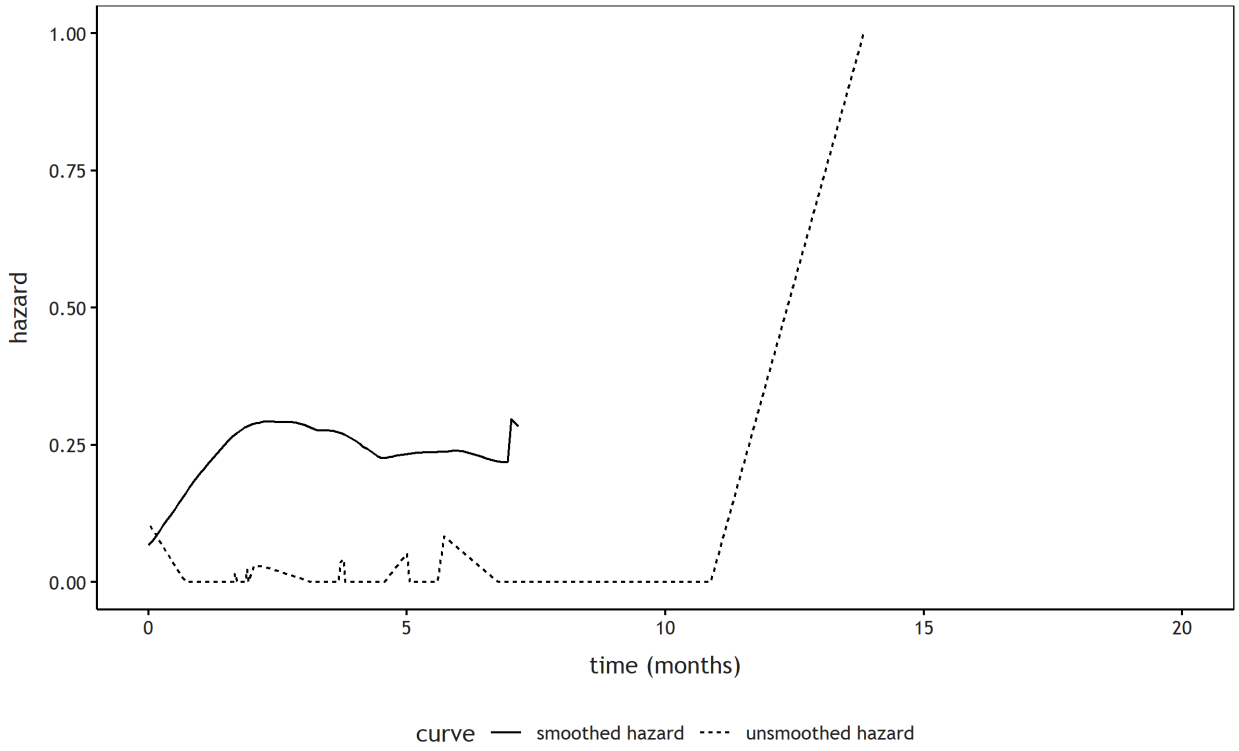


Figure 12. Modelled hazards and smoothed hazards for Placebo PFS

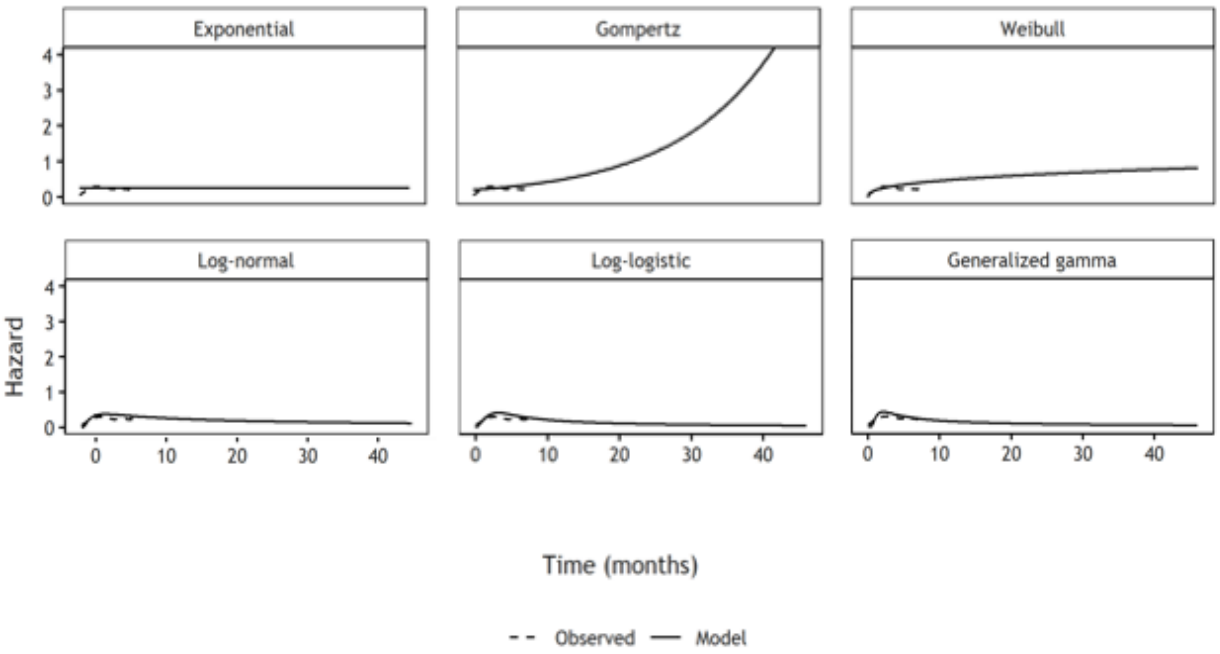


Figure 13. Smoothed and unsmoothed hazards for Cabozantinib TTD

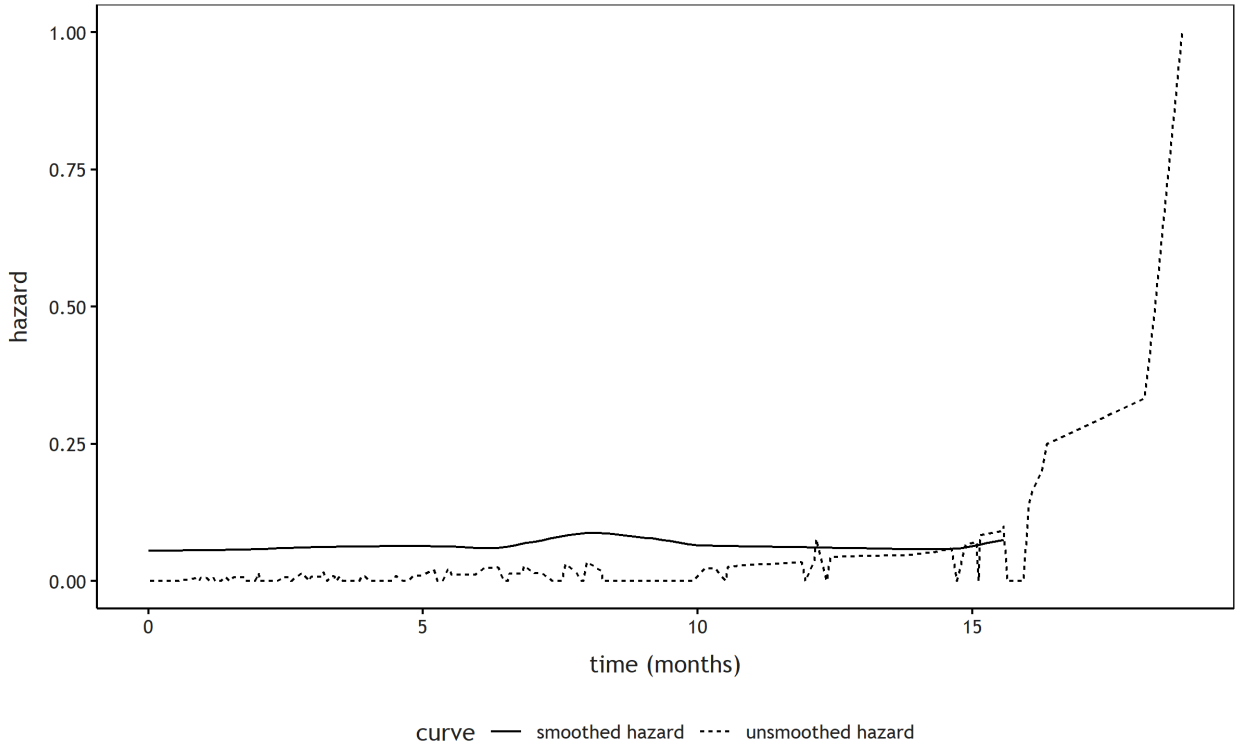
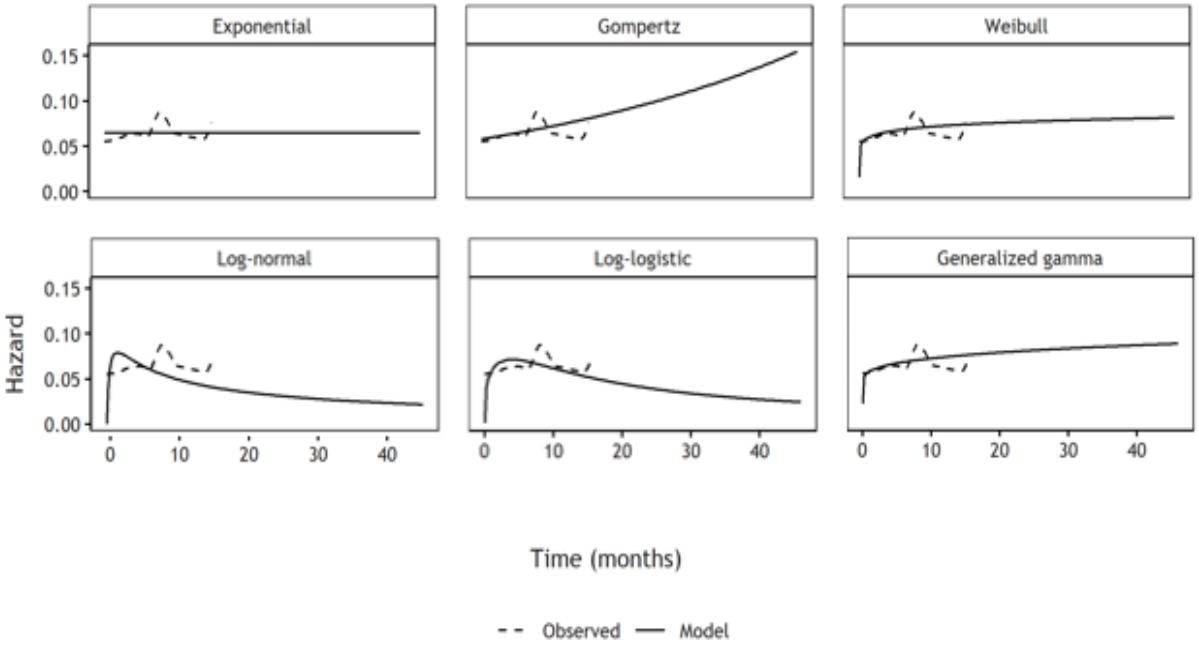


Figure 14. Modelled hazards and smoothed hazards for Cabozantinib TTD



B3. Please provide assessments for the proportional hazard assumption for PFS and OS, including log-cumulative hazard plots.

Response

The plots assessing the proportional hazard assumption between cabozantinib and placebo for OS (RPFST-adjusted) and PFS are provided below. In both cases, the log-cumulative hazards cross, suggesting that the proportional hazard assumption does not hold. The plots of the Schoenfeld residuals also suggest non-zero slopes of the scaled residuals against time. Given these results, we have rejected the proportional hazard assumption for both the OS and PFS data. Consequently, the base case of the economic model uses independent models for both OS and PFS.

Figure 15. Log-cumulative hazard plot for Cabozantinib and Placebo – OS (RPFST adjusted)

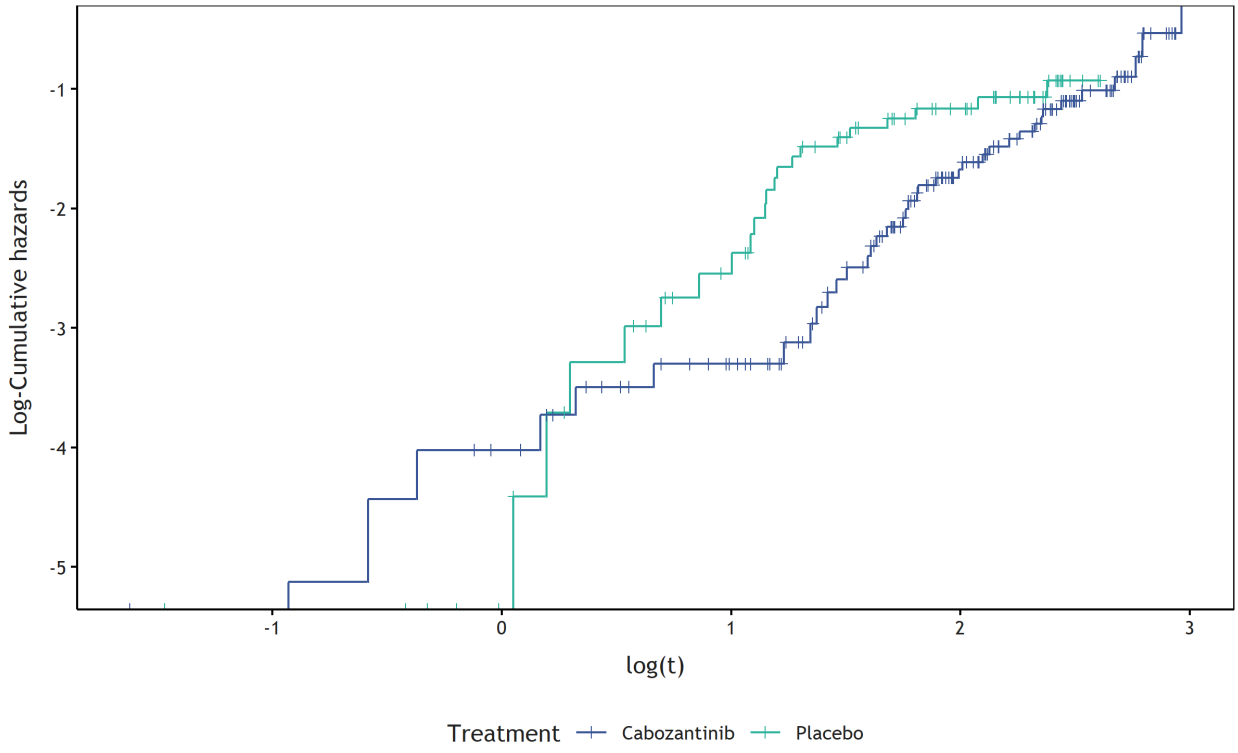


Figure 16. Plot of Schoenfeld residuals – OS (RPFST adjusted)

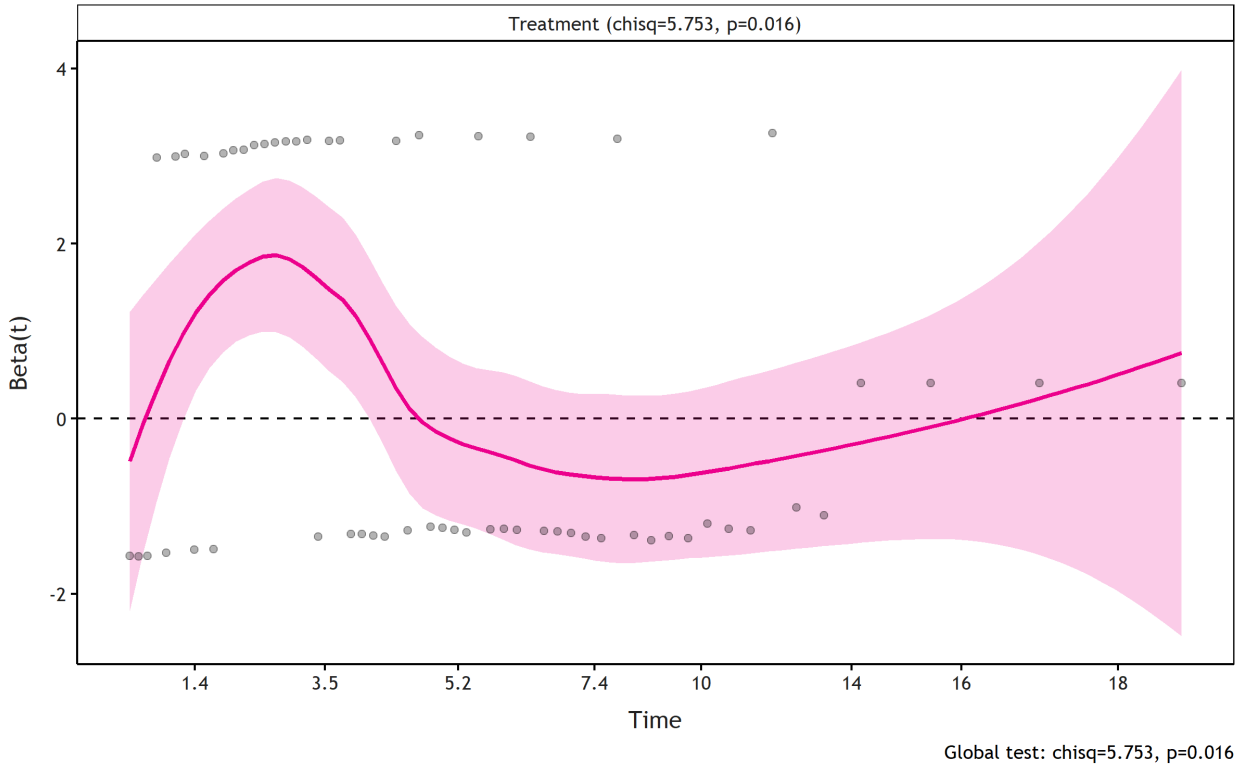


Figure 17. Log-cumulative hazard plot for Cabozantinib and Placebo - PFS

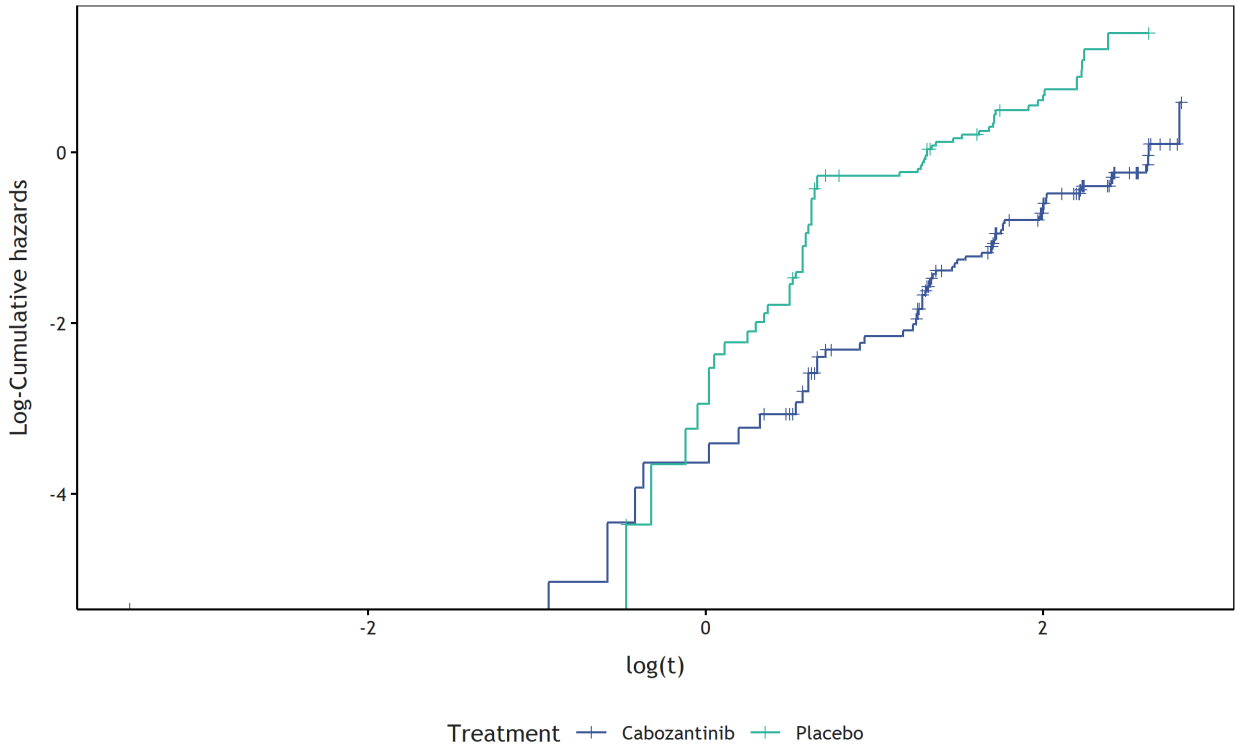
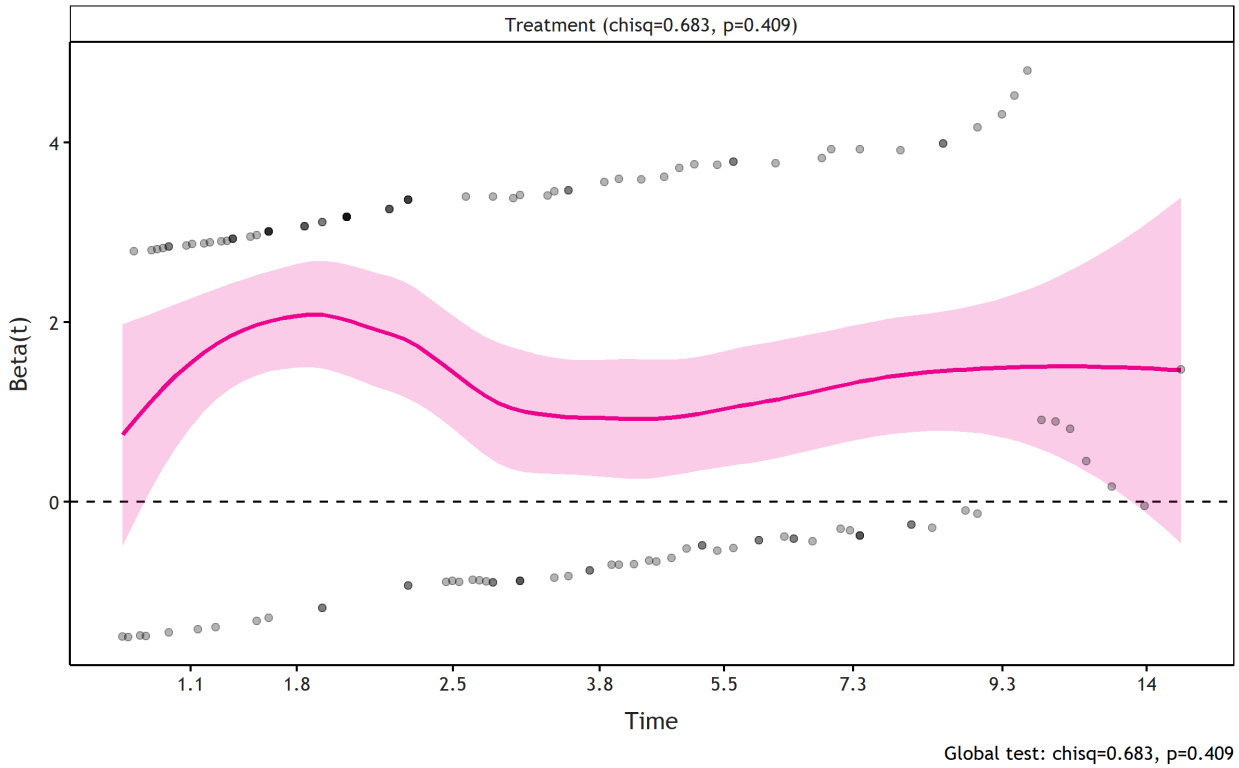


Figure 18. Plot of Schoenfeld residuals - PFS



B4. Please provide an assessment of the appropriateness of the constant acceleration factor assumption for OS in the data used (e.g., using Q-Q plots).

Response

Below we provide Q-Q plots of the observed survival data against the fitted values from an independent exponential model, as used in the base case of the economic model.

All Q-Q plots suggest that the exponential distribution, with its constant hazard, is not a very appropriate fit for the OS data from COSMIC-311 as the dots are far from a diagonal line at 45 degrees.

However, given the short follow-up time of COSMIC-311, some of the patterns conveyed by the observed data might not be reflective of what one would see in real life. An analysis of 15 years of data from the US SEER database showed that long-term survival rates for locally advanced or metastatic thyroid cancer patients are best approximated by a model featuring a constant risk of death (i.e. an exponential model). To ensure that the model assumptions are reflective of the reality of DTC, the exponential distribution was used to fit OS curves in the base case of the model, instead of relying solely on data from the COSMIC-311 trial.

Figure 19. Q-Q plot of observed survival data for Cabozantinib against fitted survival from exponential model in base case

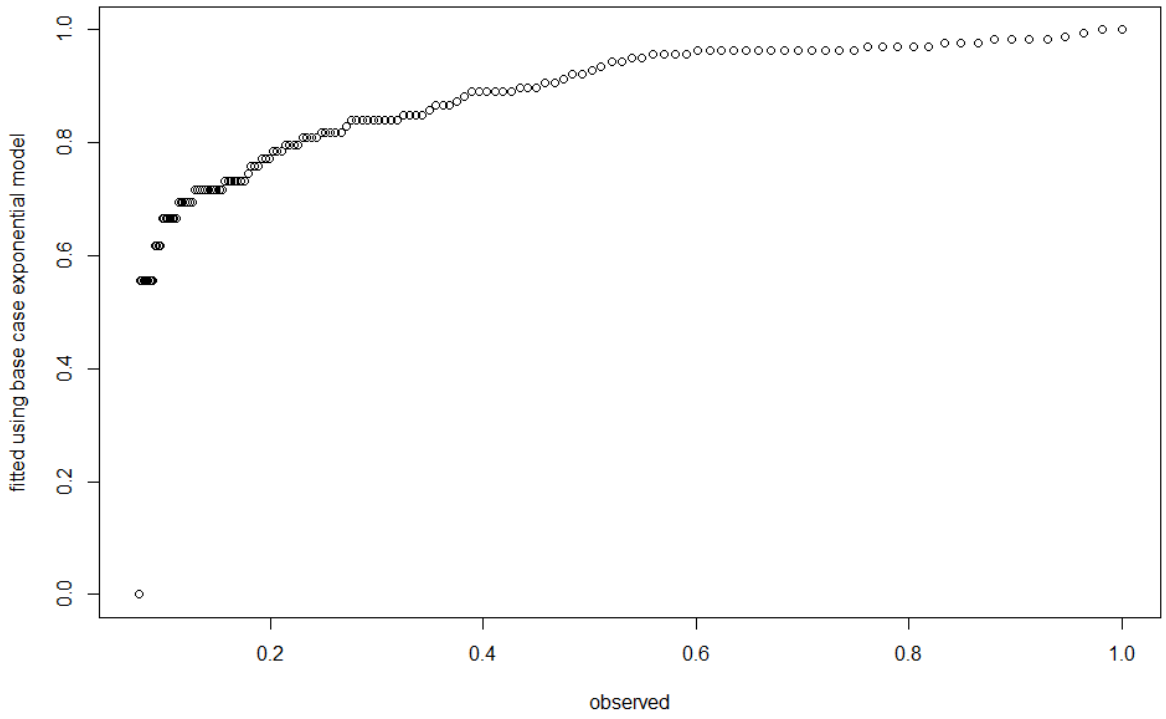


Figure 20. Q-Q plot of observed survival data for Placebo (RPFST-adjusted) against fitted survival from exponential model in base case

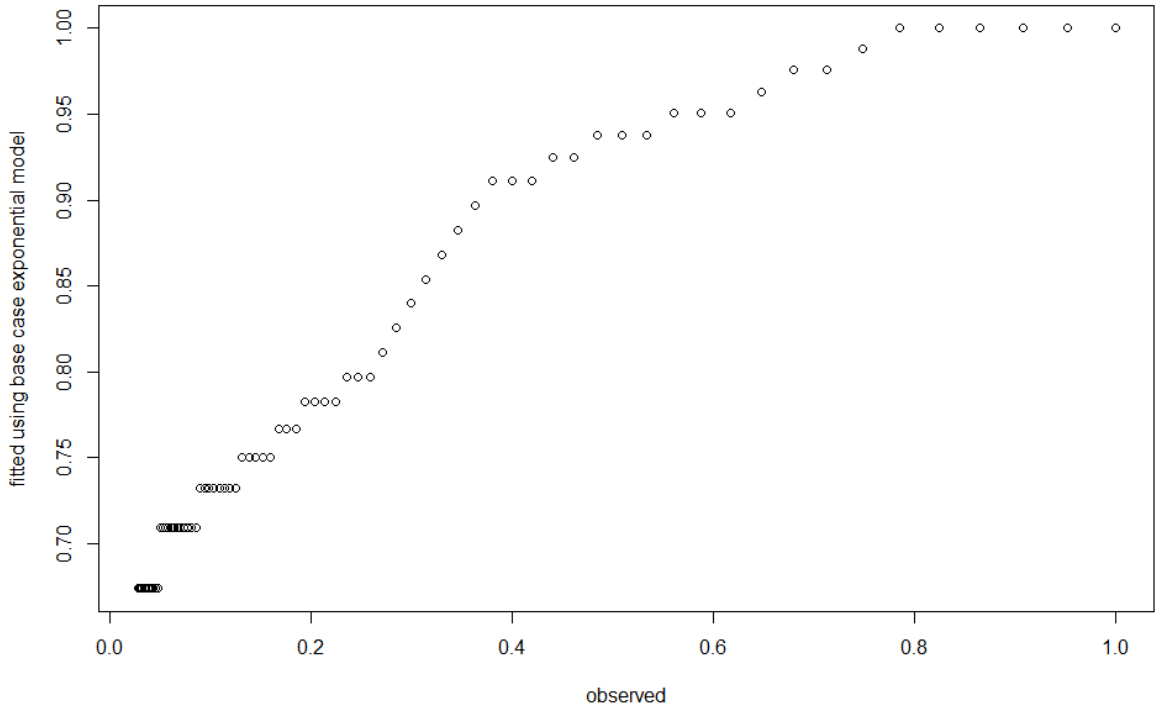
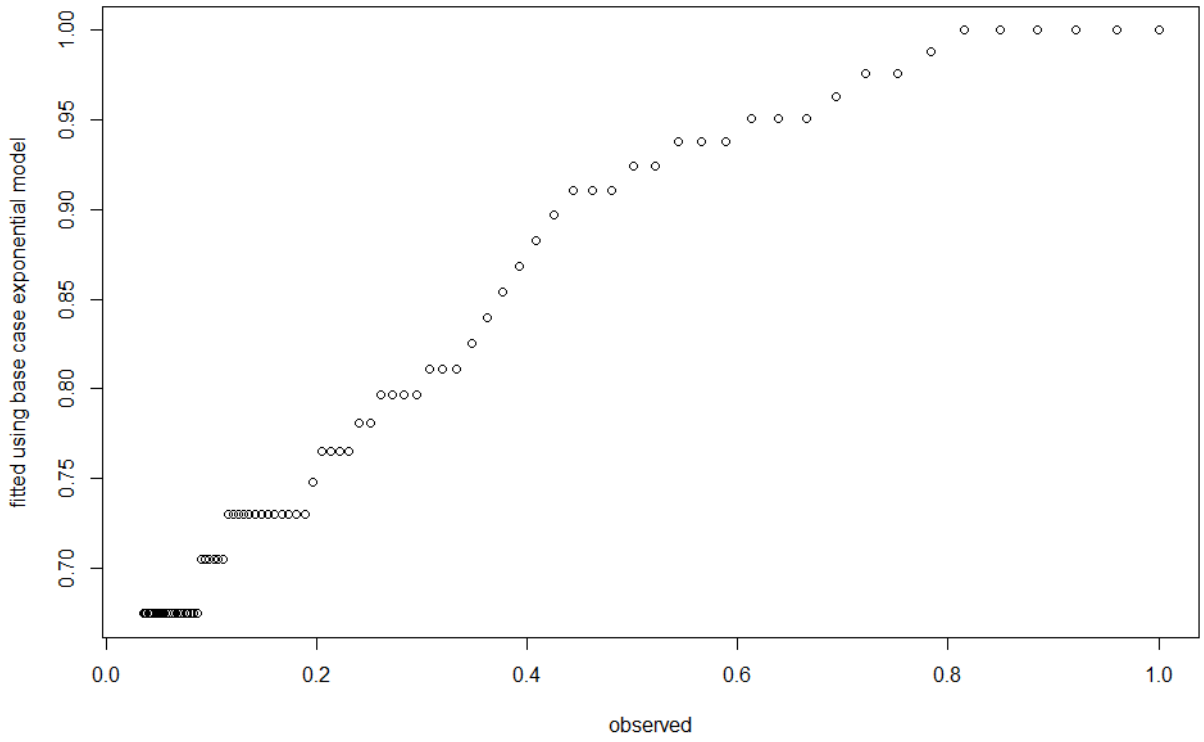


Figure 21. Q-Q plot of observed survival data for Placebo (2 stage-adjusted) against fitted survival from exponential model in base case



B5. PRIORITY. CS, Section B.3.3.4, page 105 and model, worksheets "KM Data" and "Survival Analysis". The CS states that *"Standard parametric models were also fitted to the TTD data obtained from the COSMIC-311 trial to extrapolate the TTD beyond trial duration."* However, cell O6 in model worksheet "KM Data" states *"Note: TTD data not available therefore equalised to PFS."* Despite this, the parameter values used in the survival models in worksheet "Survival Analysis" are different between PFS and TTD.

- Please clarify if TTD data are available for the cabozantinib group of COSMIC-311. If so, please provide the summary Kaplan-Meier estimates for CCO2 (in the same form as those presented for PFS and OS in model worksheet "KM Data.").
- CS Figure 27 does not clearly show the modelled TTD functions as the extrapolated functions are constrained by PFS from around 10 months onwards (hence, the tail of the function shown is the Weibull PFS model rather than the

TTD models). Please provide a comparison of the unconstrained TTD models compared against the observed TTD Kaplan-Meier estimates.

Response

TTD estimates using date of progression from CCO2 are available for the cabozantinib arm. This has been updated in the “KM data” worksheet columns K:L in the attached revised model. Figure 22 displays the extrapolated TTD functions unconstrained by PFS.

Figure 22: [redacted]

B6. PRIORITY. CS, Section B.3.3.3.2, page 98. The clinical experts who attended the 2022 Ipsen clinical advisory board meeting commented that [redacted]. The economic model applies a constraint which forces all surviving BSC patients to have died at 5 years. Please comment on whether the experts considered the OS predictions obtained from the constrained exponential model for BSC to be plausible. Also, given the experts’ concerns about overestimation of OS in both groups, please explain why no adjustment or constraint was made to the cabozantinib group in the model.

Response

As referenced in the response to B1, the model now incorporates CCO1 OS for the BSC arm as part of the base case, removing the need for the 5-year OS constraint

applied in the original model. The decline in the observed trial CCO1 OS data is more reflective of that seen in clinical practice without a flattening in the tail such that the extrapolated curve has a smooth decline and no longer has a sharp drop in OS at year 5, as was initially programmed in the model.

The standard parametric curves for OS for BSC following RPFST-adjustment are presented in Figure 23.

Table 3 presents the AIC and BIC values for the models. According to the AIC and BIC, the Weibull was the best fitting model for cabozantinib and Generalized gamma for BSC adjusted using RPSFT.

Table 4**Error! Reference source not found.** displays the survival estimates for cabozantinib and BSC at different timepoints and distributions. Log-logistic and Log-normal curves were deemed not plausible as both distributions overestimated survival aligned with the UK clinician opinion provided in the advisory board used to validate the original curve selection.[13] Additionally, the Generalized gamma produces a BSC curve which crosses with the cabozantinib OS curve and is therefore not considered plausible. The remaining distributions (Exponential, Weibull and Gompertz) produce more plausible estimates, with the Exponential distribution having one of the best statistical fit coupled with the best visual and clinically plausible fit and has therefore been selected as the most appropriate distribution for the model base case (Figure 24) with updated base case results in the Appendix. A scenario where Weibull, the

second best curve in terms of visual and statistical fit, has also been included in the Appendix.

Figure 23: [redacted]

[redacted]

Table 3: [redacted]

[redacted]	[redacted]			[redacted]			[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted] [redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

* This distribution was selected as the best fitting model, based on minimisation of AIC and visual fit.
Abbreviations: AIC – Akaike's Information Criterion; BIC – Bayesian Information Criterion; BSC – Best supportive care; RPFST – Rank Preserving Structural Failure Time.

Table 4: Proportion of individuals alive in the cabozantinib and BSC arms

Abbreviations: BSC – Best supportive care; Cabo – Cabozantinib; PFS – Progression-free survival
* This distribution was selected as the best fitting model, based on minimisation of AIC/BIC, visual assessment, and clinician validation, with which to extrapolate the COSMIC-311 OS data – CCO1 is Full ITT population with RPSFT adjustment.

Figure 24:

Cabozantinib remains unconstrained in the economic model base case. A 2017 study explored the use of cabozantinib (60 mg) as salvage therapy for RAI-refractory DTC patients.[14] The median duration of follow up was 22.8 months (95% CI 21.2 - 30.2) and median OS was reported as 34.7 months (95% CI 18.3 - not reached). Comparatively, the COSMIC-311 CCO2 follow was less at 10.1 months, and median OS was smaller at 19.4 months.[14] The company have reached out to the authors of the 2017 study to explore the option of incorporating their study results into the economic model. Whilst we wait, the CCO2 observed data for cabozantinib OS conservatively remains the base case for the model. A scenario which uses CCO1 data for cabozantinib and BSC has been provided in the Appendix.

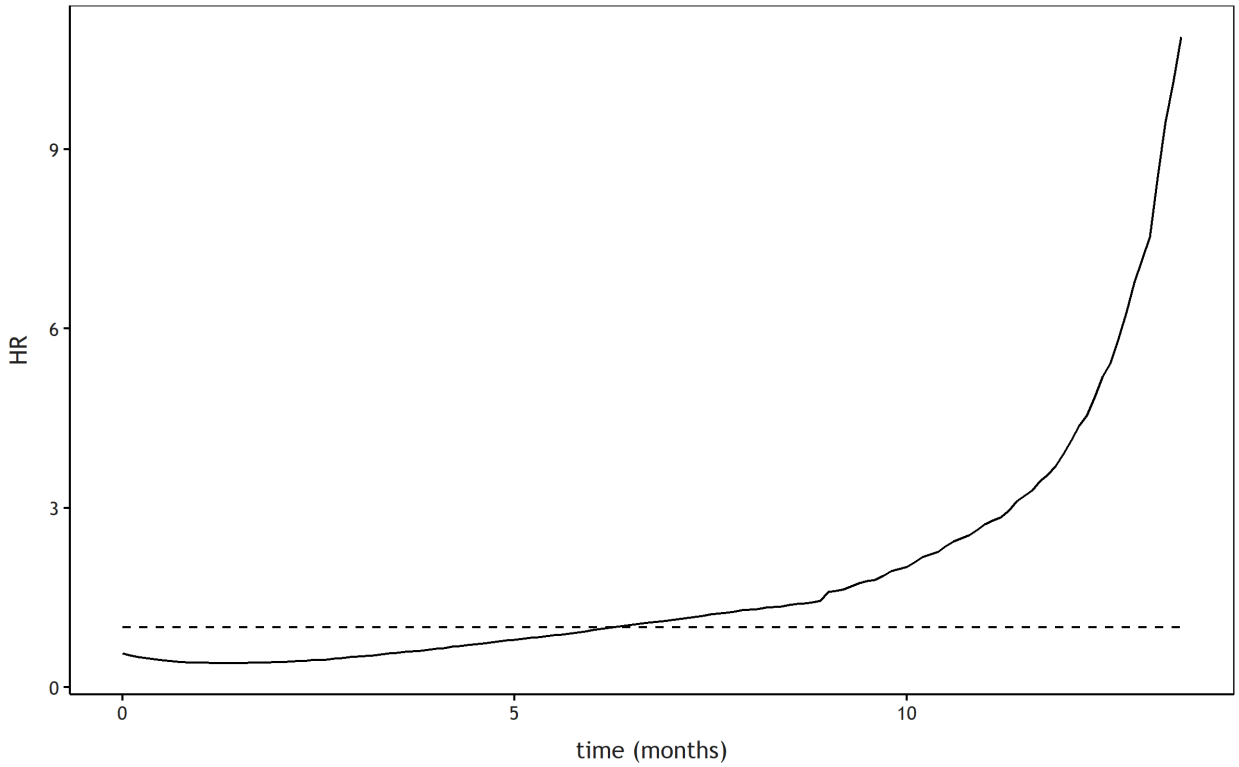
B7. PRIORITY. CS, Section B.3.2.2, Table 20, page 89. The model applies exponential distributions to OS in both groups, with a lower hazard applied in the cabozantinib group. This assumes an indefinite relative treatment effect. Please justify this assumption. Please also provide a plot of the time-varying HR for the observed cabozantinib and the RPSFTM-adjusted placebo OS data.

Response

The use of independent exponential curves to fit OS in both treatment arms was motivated by the analysis of 15 years of data from the US SEER database, which showed that long-term survival rates for locally advanced or metastatic thyroid cancer patients are best approximated by a model featuring a constant risk of death (i.e. an exponential model). This approach (using an exponential model to extrapolate OS) was used for differentiated thyroid cancer in NICE Technology Appraisal (TA535) for lenvatinib and sorafenib.

Below we provide a plot of the observed (smoothed) HR between cabozantinib and RPFST-adjusted placebo over time. Note that, given the short follow-up time of COSMIC-311, some of the patterns shown in the data are not considered clinically plausible, such as a HR of cabozantinib versus placebo for OS that is greater than 1. This strengthens the argument for using external data sources, such as the SEER database, to guide the long-term extrapolations, as opposed to relying solely on data from the COSMIC-311 trial.

Figure 25. Observed HR for OS of Cabozantinib vs Placebo (RPFST adjusted) over time



Notes: dashed horizontal line represents HR = 1.

B8. Please provide the R code used for the IPCW, two-stage and RPSFT methods.

Response

[redacted]

[redacted]

[redacted]

B9. CS, Section B.2.6.4, pages 51-54 and Section B.3.3.3.2, pages 96-97. Please clarify whether re-censoring was applied to both the two-stage and RPSFT analyses presented in the clinical and economic sections of the CS. Please provide justification for the chosen option.

Response

Re-censoring was considered for both the two-stage and RPFST models in order to account for the maximum survival observation in the trial. Consequently, the counterfactual survival data estimated by the RPFST model does not extend the survival observed in the trial.

For the RPSFT method, the amount of re-censoring was limited. The survival of N=8 patients was re-censored. The survival time based on the re-censored data was shorter (median: 1.0 months; min – max: 0.14-3.7 months). However, due to the low level of re-censoring and the differences in estimated survival, the effect of re-censoring remains limited.

For the two-stage method, re-censoring was considered. However, the estimated counterfactual data via the two-stage method did not exceed the maximum observed survival. Therefore, no re-censoring was applied.

B10. CS, Section B.2.4.2.2, page 37. Please clarify which baseline and time-dependent characteristics were adjusted in the IPCW analysis. Please also clarify how these covariates were identified and selected.

Response

In the IPCW method, the probability of cross-over was estimated as a function of time (modelled as a quadratic effect), a time-dependent progression variable (flagging patients who progressed within the next 34 days, which was the median time from progression to switch to cabozantinib observed in the trial), age group and previous use of lenvatinib. Consequently, the IPCW method accounted for the variables that were most important clinically which were also used in the RCT for treatment stratification, the baseline variables age group ($\leq$ 65 & >65) and prior lenvatinib use.

The sample size of eligible patients is limited, only N=88 patients received placebo of which N=65 patients were eligible to switch, i.e. progressed (of which N=39 switched

treatment and N=26 did not switch). Furthermore, other potentially important variables, such as patient preference for switching or time-dependent variables were scarce and consequently could not be adjusted for. Based on the limited sample size only the most clinically relevant variables were included, i.e. the variables used were the variables also used for stratification in the clinical trial.

B11. CS, Section B.2.4.2.2, page 38. Regarding the two-stage method for adjusting for direct treatment switching:

- Please clarify which covariates were included in the AFT model for the two-stage method. Please also clarify how these covariates were identified and selected.

Response

  - Variables included in the two-stage AFT model were age group (≤ 65 & >65) and prior lenvatinib use.
  - The sample size of eligible patients is limited, i.e. the treatment effect calculation is based on N=65 patients in total that received placebo and progressed (N=39 switched treatment and N=26 did not switch). Based on the limited sample size only the most clinically relevant variables have been assessed to prevent overfitting, i.e. the variables used were the variables also used for stratification in the clinical trial.

- The AIC for the log-logistic and log-normal models in the two-stage method were similar (i.e., within 3 points difference). Please provide the point estimate for the HR and 95% CI for cabozantinib versus placebo-two-stage adjusted using the log-logistic model.

Response

  - HR point estimates and 95% CI only changed slightly when using the log-logistic model:

	Cabozantinib	Placebo	Placebo-adjusted
Events, n (%)	37 (22)	21 (24)	21 (24)
Median OS (95% CI)	19.35 (15.87, NA)	NA (NA, NA)	NA (NA, NA)
Unstrat HR	NA	0.78 (0.45, 1.33)	0.74 (0.43, 1.27)
Strat HR	NA	0.76 (0.45, 1.31)	0.71 (0.41, 1.23)

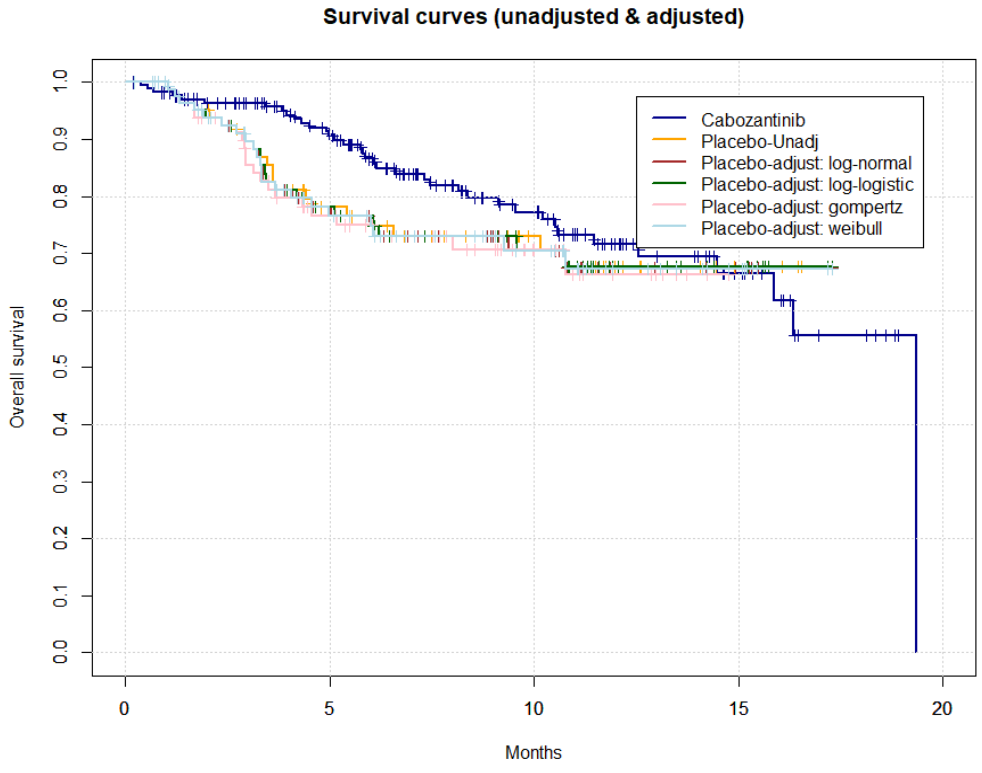
- Please provide the results using the gamma or generalised gamma model in the two-stage method. Please include AIC, visual inspection assessment, the point estimate for the HR and 95% CI for cabozantinib versus placebo-two-stage adjusted.

Response

  - The current two-stage adjustment has been conducted using the R package 'eha' (https://cran.r-project.org/web/packages/eha/eha.pdf). While this package allows for the assessment of different models, it does not include the functionality to assess the two-stage method using the gamma or generalized gamma model. The assessment of these two additional distributions with the same R package and model is therefore not possible and as such is left out for this analysis. In total four different distributions were examined producing mostly highly overlapping survival estimates indicating the robustness of the estimates.

- Please provide a plot containing the fitted AFT models and the observed Kaplan-Meier function in the two-stage method.

Response



- Please comment on the clinical plausibility of the projections for all AFT models used.

Response

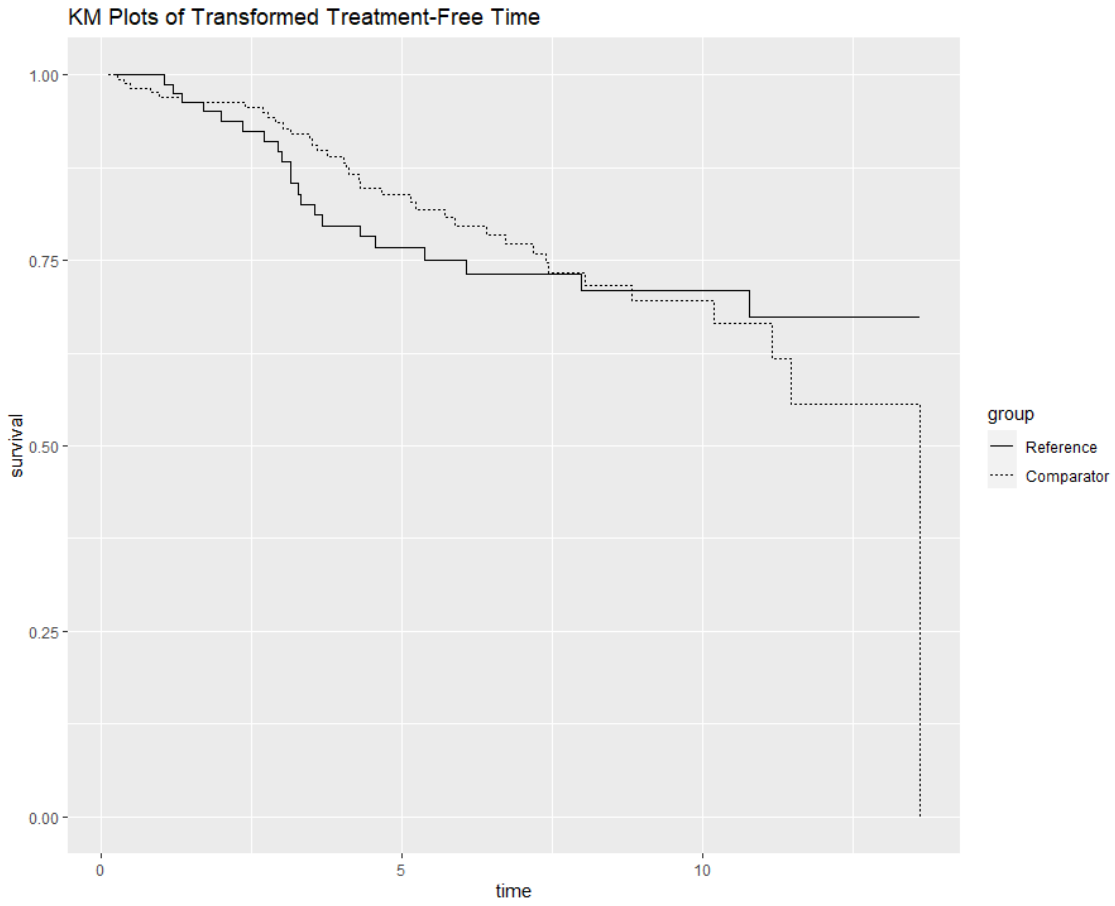
- All models used to adjust the survival in the placebo arm via the two-stage method produced similar adjusted survival estimates indicating the robustness of the analyses. The estimated treatment effect of the three best fitting models was 9.3%, 6.7%, and 10.7% while the estimated treatment effect with the gompertz model was considerably higher with 31%. However, the impact of the model choice for the two-stage method on the estimated survival appears to be very limited with the Placebo-adjust: log-normal, Placebo-adjust: log-logistic, Placebo-adjust: Weibull curves mostly overlapping. Also, the considerably higher and potentially overestimated treatment effect of the gompertz model does seem to produce limited differences in the estimated survival.

B12. CS, Section B.2.4.2.2, page 37. Regarding the RPSFT method for adjusting for direct treatment switching:

- Please provide a plot comparing the untreated survival curve for the BSC group and the untreated survival curve for the cabozantinib.

Response

- The RPFST method estimates the treatment effect (psi) by balancing counterfactual event times of the two treatment arms. Counterfactual event times represent event times that would have been observed if no treatment were received. At the estimated psi, the plot below presents the counterfactual KM curves of the groups overlapping and crossing, i.e. the distributions of counterfactual event times at the estimated psi are similar and fulfilling this assumption.



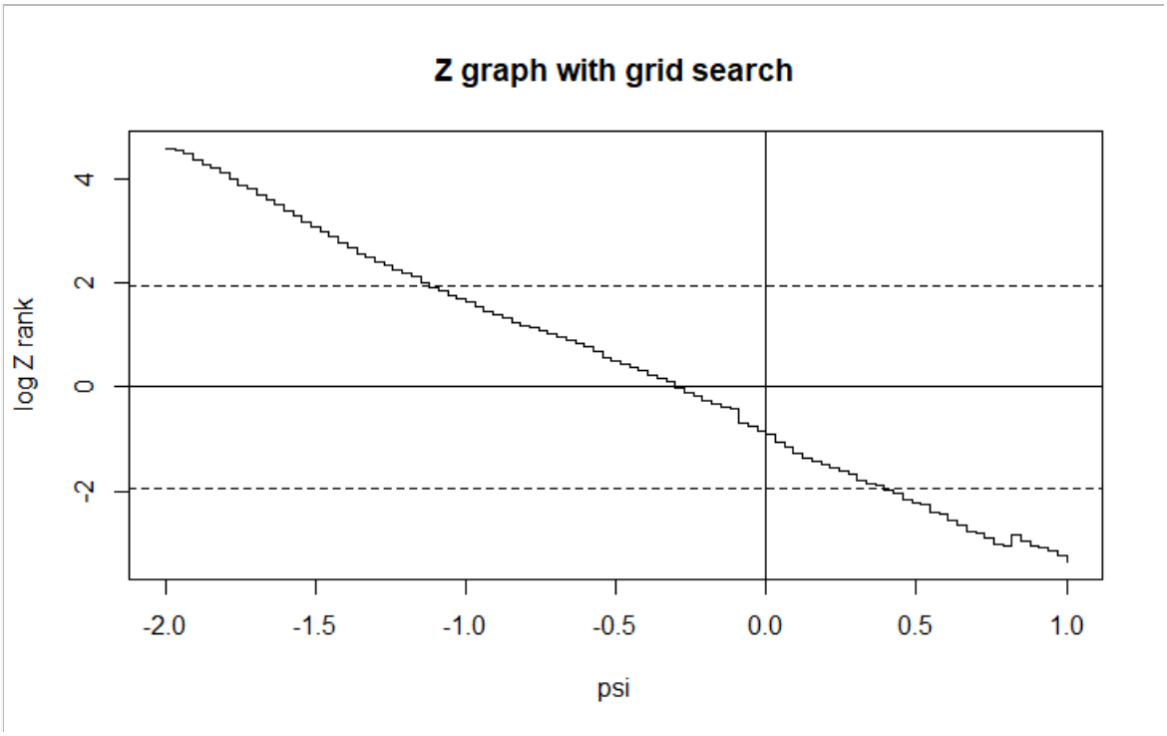
Reference group = placebo

Comparator = cabozantinib

- Please provide the “z graph” for the g-estimation process.

Response

- The z-graph below present the test statistic Z(psi) vs. psi. The best estimate of psi is reached at Z(psi)=0, with psi= -0.3529268 (95% CI: -1.1336046, 0.3864174). The plot does not indicate any problems in the estimation of psi.



- Please clarify whether the treatment effect of cabozantinib over placebo reported in the treatment crossover report Section 3.1.1 (reported value = [redacted]) represents an acceleration factor.

Response

- Yes, this is correct. It represents an acceleration factor and the estimated effect of cabozantinib is to extend life by [redacted].

B13. PRIORITY. Model, worksheet "Survival Analysis". Section B.2.6.4 of the CS reports the results of IPCW, RPSFTM and the two-stage method for adjusting for direct treatment switching in the placebo group. However, the executable model only includes parametric survival models fitted to the OS data adjusted using RPSFTM with re-censoring. Please explain why other methods have not been explored in sensitivity analyses. Please include additional functionality in the model to explore all methods with and without re-censoring (where applicable).

Response

In the NICE appraisal document,[15] the committee agreed that the RPSFT method was the most appropriate to adjust for the high level of crossover in the SELECT[16] (88% of placebo patients crossover to lenvatinib) and in DECISION[17] (75% of placebo patients crossover to sorafenib) trials after disease progression. Given the high percentage

(45%) of placebo patients who crossed over into the active treatment arm upon progression, a RPSFT method was selected for the base case. In addition, response to B14 provides results for sensitivity analyses which supports the conclusion that the RPFST methodology is the most appropriate cross-over adjustment method for COSMIC-311.

A limitation of the IPCW method is that it relies on the “no unmeasured confounders” assumption, i.e., data must be available on all baseline and time-dependent prognostic factors for mortality that independently predict switching. This also includes key predictors of treatment switching which are often not collected in trials (e.g., patient preference for switching). Furthermore, models predicting the switching risk must be accurately specified.[18] One advantage of the IPCW method over the two-stage method is the IPCW adjusts for any differences in patient characteristics that occur between the time point of the secondary baseline, i.e., disease progression, and the time of treatment switch (e.g., laboratory values).

However, as limited covariate data were collected after progression within the trial, potential time-dependent confounding occurring between time of treatment discontinuation/progression and the time of treatment switch could not be adjusted for by the IPCW model; as such providing no advantage over the two-stage method. Among the 88 patients in the placebo arm there were only 21 deaths in the CCO2 dataset. Additionally, bias associated with the IPCW method could be high in scenarios in which the proportion of placebo group patients who switched is high (45%), leaving very few patients who didn’t switch. The IPCW method is not stable when the proportion of switchers is large and there are small sample sizes ([redacted]).

Furthermore, the COSMIC-311 study was not powered for OS and was not planned to collect information on all baseline and time-varying characteristics that are prognostic for survival. In addition, due to the limited sample size the variables adjusted for were age group and prior lenvatinib use only, aligning with the trial stratification factors, and among the 21 deaths in the CCO2 dataset there were very few deaths events which could be used (13 events). Therefore, the IPCW model could

not provide an accurate picture of the impact of treatment switching on survival and the assumption of “no unmeasured confounding” underlying the IPCW may be violated.

A previous study found that IPCW is prone to bias in small samples, if selection bias was very strong, and if there were unmeasured confounders.[19] As a result, the IPCW method has not been implemented in the model.
A limitation of the two-stage model is the assumption that there is no time-dependent confounding between the time of disease progression and the time of treatment switch. As only stratification factors used within the COSMIC-311 trial (i.e., age group and prior levantinib use) were used as covariates the two-stage method applied to COSMIC-311 may also be subject to residual confounding, so was not selected for the base case analysis. However, a scenario using the two-stage method has been added to the CEM with results presented in the Appendix.

B14. PRIORITY. Company’s treatment crossover report, Section 3.1.1, page 15. The treatment crossover report contains sensitivity analyses which vary the value of *k* in the RPSFTM analysis. These have not been included in the economic model. Please include additional functionality in the model to allow the user to perform these sensitivity analyses.

Response

Due to the time-intensive nature of this request and short window for providing our responses, the Company have not added this functionality into the model. In the absence of this, we have provided the results of the sensitivity analyses below which support the conclusion that the RPSFT methodology is the most appropriate cross-over adjustment method for COSMIC-311.

The sensitivity analysis assessed “common treatment effect assumption” where different treatment effect assumptions were used. The treatment effect parameter k ranged from 0 to 1 with k=0 assuming no treatment effect in crossover patients and k=1 assumes the common treatment effect. The results from those sensitivity analyses are presented in Table 5.

Table 5: Stratified HR (inflated 95% CI) for cabozantinib vs placebo-RPSFT adjusted per different treatment effect assumptions (k)

Abbreviations: CI – Confidence interval; HR – Hazard ratio; NA – Not applicable; RPSFT – Rank-preserving structural failure time

The sensitivity analyses found that the point estimate of the hazard ratio (HR) between cabozantinib and placebo-RPSFT adjusted varied consistently between , indicating it has a relatively small impact on the overall estimated relative treatment effect of cabozantinib vs placebo after adjustment. In turn application of these sensitivity analyses in the model would be expected to have a minimal impact on the results, furthermore assuming a common treatment effect assumption is appropriate such that RPFST is an appropriate cross-over adjustment method.

B15. PRIORITY. CS, Section B.2.6.4, page 53. The CS states *"Although the RPSFT and the two-stage results are more likely most appropriate for adjusting treatment crossover in the COSMIC-311 trial, the RPSFT method has been used as the base case because it was in line with previous NICE submissions, in particular TA535."* Given that the time between progression and switching was short, please clarify why the two-stage method was not preferred over RPSFTM.

Response

As indicated in B13, given the high percentage (45%) of placebo patients who crossed over into the active treatment arm upon progression, the unadjusted placebo OS curve is confounded, as it does not account for the potential treatment benefit received by those patients that crossed over to the active treatment. The RPSFTM method was preferred within the base case as it accounted for this bias as it assumes that treatment effect is the same regardless of when the experimental treatment is initiated.
Regarding the 'two-stage' method, numerous 'subjective' factors influence which patients are selected to cross over.[20] Therefore, it is difficult to justify the 'no unmeasured confounders assumptions' as it requires all covariates and time-dependent factors determining cross-over to be known and measured at appropriate time points in the trial. Due to limited sample sizes, only stratification factors were used as covariates in the analysis (i.e., age group and prior lenvatinib use), therefore the major assumption required for the two-stage method to remain valid is unlikely to hold. However, as stated in B13, the two-stage method has now been implemented within the model for transparency with results presented in the Appendix.

B16. Model, worksheet "Survival Analysis". Please explain how additional uncertainty associated with artificial censoring in the RPSFTM analysis has been included in the probabilistic sensitivity analysis used in the executable model.

Response

No additional uncertainty associated with artificial censoring in the RPSFTM analyses has been included in the PSA, other than the uncertainty tested by the PSA itself which explores the impact of model parameters uncertainty on the results.

Executable model

B17. The EAG has identified five errors in the executable model which are listed below. Please explore these issues, confirm that they are errors and provide a revised version of the executable model.

(a) Model, worksheet "Survival Analysis", cells C64:W663. The half-cycle correction is applied incorrectly as it overestimates the contribution of the first cycle to overall health outcomes and costs. The half-cycle correction should be applied by taking the average of the modelled cumulative survival

probabilities between consecutive cycles for each endpoint (PFS, OS and TTD).

(b) Model, worksheet “Data Store”, cells D122:E223. The values used in this cell range are arbitrary numbers which increase in increments of 0.001 in each year. Life tables for England should be used.

(c) Model, worksheet “Clinical Inputs”, cells N67:N667. The general population mortality risks assume that the same proportionate split of men and women are alive in each cycle. However, ONS life tables show that men and women have different annual risks of death by age. In addition, the constraint in model worksheet “Clinical Inputs” columns F and G is applied only to the cumulative OS probabilities, rather than the per-cycle risk of death. The EAG prefers an approach which assumes the sex distribution in COSMIC-311 applies at time zero and estimates the cumulative survival probabilities using annual life table probabilities for each age (i.e., a weighted survival model). The constraint should be applied to the risk of death in each cycle, not to the cumulative survival probabilities.

(d) CS, Section B.2.10.2, pages 70-71 and model worksheet “Quality of Life Inputs”, C20:K39. The AE frequencies have been inappropriately rounded down to integer values.

(e) Model, worksheet “Trace (Cabozantinib)” and “Trace (BSC)”, columns AH and AI. The model applies age-adjusted utility values which are higher than the average EQ-5D in the general population. A cap should be included.

Response

We acknowledge the EAG’s list of errors and can confirm the following corrections have been made to the model:

(a) The half cycle model correction has been updated as per the EAGs request. The updated formulae can be found in Columns E9:G608, and Column I9:I608, in both the “Trace (Cabozantinib)” and “Trace (BSC)” worksheets.

(b) The life tables for England[21], based on data for the years 2018-2020, have been incorporated into the model. This can be found in the "Data Store" worksheet, cells D228:E329.

(c) The weighted survival model approach was originally programmed into the model. Upon updating the life tables with male and female specific mortality data, the weighted survival probabilities have automatically updated to reflect the sex distribution in the COSMIC-311 trial. Additionally, the Company have updated the constraint in the "Clinical Inputs" worksheet to the EAGs preferred approach of risk of death per cycle. This can be found in cells F67:G667 in the "Clinical Inputs" worksheet, using general population risk of death per cycle from P67:P667, and treatment specific risk of death per cycle in Y64:Y663 and AX64:Y663 in the "Survival Analysis" worksheet for cabozantinib and BSC, respectively.

(d) AE frequencies have been updated in "Data Store" F197:F208 to reflect the data presented in Table 18 in the CS.

(e) An age-adjusted general population utility cap has been applied to the PFS and PD heath states for both cabozantinib and BSC. This can be found in cells AH9:AI608 in "Trace (Cabozantinib)" and "Trace (BSC)". Age-adjusted general population utility values have been sourced from Hernández Alava et al. (2022).[22]

HRQoL

B18. CS, Section B.3.4.1, pages 106-109. With respect to the utility model fitted to EQ-5D data from COSMIC-311:

- Please clarify which covariates (e.g., age, gender, treatment arm etc.) were included in the preferred model for the utility analysis.
- Please clarify how the preferred model was chosen.
- Please provide the results of the full utility model including all covariates considered in the same model.
- Please clarify which data-cut was used for the utility model. If CCO2 was not used, please clarify why this is the case.

Response

Linear mixed-effect models were used to derive health state utility values ranging from 0 to 1. Several model structures were considered, including random intercepts, random slopes, and random intercepts and slopes. Several potential covariates were included in the models, such as age, gender, treatment arm, assessment time points, and progression state. The preferred model structure included a random intercept at the subject level.

$$Y_{it} = \beta_0 + \beta_1 hs_{it} + u_i + e_{it}$$

Where Y_{it} denotes the EQ-5D-5L utility value measured for patient i at time t, e_{it} is the random error term, and u_i is the random intercept term.

Several considerations were taken into account when choosing the base-case utility model, including statistical model fit, sample size and the requirements/capabilities of the cost-effectiveness model. The models tested, their covariates and AIC and BIC fit results are displayed in Table 6. The best fit in terms of AIC and BIC was Model 1, including only the health state covariate. This reflects the cost-effectiveness models capabilities as states for response and treatment status have not been included. It also reflects the use of health state utility values pooled across treatment arms with AEs captured separately, as it is unlikely that AEs are fully captured on the EQ-5D assessment visits within the recall period of "today" and any AEs that are captured are implicitly assumed to apply for the full duration of time between HRQoL assessments, regardless of the true duration of the AE in practice. Therefore, Model 1 was selected as the recommended model for use in the cost-effectiveness model.

Table 6: Comparisons of model fit

	Covariates included	AIC	BIC
Model 1	Health state	[redacted]	[redacted]
Model 2	Health state + response + treatment arm + treatment status + end-of-life	[redacted]	[redacted]
Model 3	Health state + treatment arm	[redacted]	[redacted]
Model 4	Health state + treatment status	[redacted]	[redacted]
Model 5	Health state + response	[redacted]	[redacted]

Abbreviations: AIC – Akaike information criterion; BIC – Bayesian information

Coefficients, the variance covariance matrix and the resulting utility values for Model 1 are displayed in Tables 5-7. In Model 1 the utility in the PFS state is [redacted] (Table 7). Utility in the PD state is [redacted] lower, at [redacted].

Table 7: Utility Model 1 coefficients

Parameters	Estimate	SE	p-value	95% Confidence interval of estimate	AIC	BIC
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Health state: PD	[redacted]	[redacted]	[redacted]	[redacted]		

Abbreviations: AIC – Akaike information criterion; BIC – Bayesian information; PD – Progressed disease; SE – Standard error

Table 8: Utility Model 1 Variance Covariance Matrix

	Intercept	Health state: PD
Intercept	[redacted]	[redacted]
Health state: PD	[redacted]	[redacted]

Abbreviations: PD – Progressed disease

Table 9: Utility Model 1 utility values per health state

Parameters	Health state value
Progression free	[redacted]
Progressed disease	[redacted]

A response for *"Please clarify which data-cut was used for the utility model. If CCO2 was not used, please clarify why this is the case"* will shared on the 26th October.

B19. Model, worksheet "Data Store", cells C226:I278. The model includes age-adjustment of health state utility values based on the regression equation reported by Ara and Brazier (2010) and utility values from Fordham *et al*. (2015). These calculations appear to calculate age-adjusted utility multipliers by assuming a *"source publication population age"* of 67 years. However, the Fordham paper does not report a mean age of 67 years and a newer set of general population EQ-5D weights for the UK has been reported by Hernandez Alava *et al*. (2022). Please amend the age-adjustment of utility values in the executable model by calculating age-adjusted utility value multipliers using the Hernandez Alava *et al*. EQ-5D dataset assuming a mean age of 65 years.

Response

As per the response to B17(e), age-adjusted utility value multipliers have been updated using the Hernandez Alava *et al.* (2022)[22] EQ-5D data set. This is located in the ”Data Store” worksheet, cells J340:L426. The updated model and appendix with results will be shared on the 26th October.

B20. PRIORITY. CS, Section B.3.4.1, page 106. The utility value for the progression-free state in Fordham *et al.* (2015) is substantially higher than that obtained from the analysis of EQ-5D data in COSMIC-311. Please justify why the data from COSMIC-311 have not been used to inform the utility value for the progression-free health state in the model.

Response

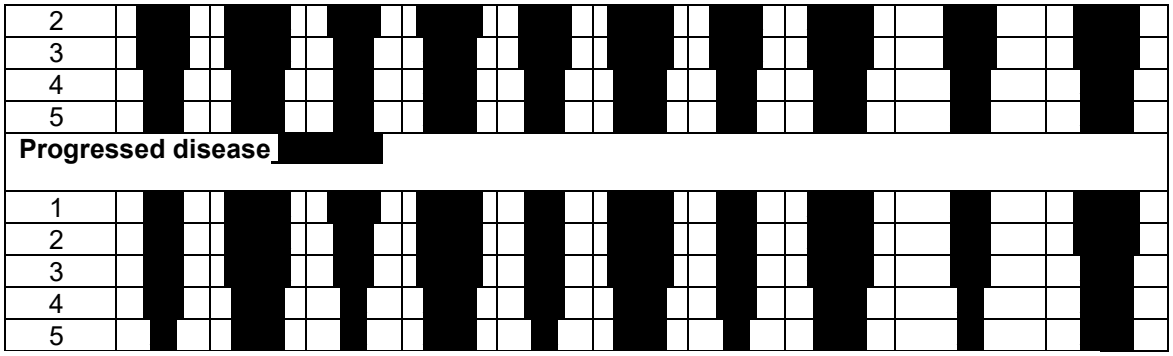
The EQ-5D-5L data collected within the COSMIC-311 trial[23] was analysed to estimate health state utility values. In the COSMIC-311 trial, EQ-5D responses were provided by patients at various assessment/time points. For those patients who crossed over treatment, utility assessments were discontinued. As a result, it was not possible to obtain post-crossover specific utility values.

The EQ-5D-5L data from the COSMIC-311 trial was mapped to the EQ-5D-3L using the cross-walk approach by Hernandez-Alava and Pudney (2017)[24] as dictated by recently published NICE guidelines (2022)[25]. The health state utility values from the COSMIC-311 analysis are [redacted] for PFS and [redacted] for PD. In addition, the utility value of patients in the death state is assumed to be zero, as per standard convention.

The limited impact of progression is unsurprising given the small differences in EQ-5D-5L responses between health states as seen in Table 10.

Table 10. Response distribution counts and percentages in utility sample ([redacted])

Level	Mobility		Self-care		Usual activities		Pain/ discomfort		Anxiety/ depression	
	N*	%	N*	%	N*	%	N*	%	N*	%
Full utility sample [redacted]										
1	[redacted]									
2	[redacted]									
3	[redacted]									
4	[redacted]									
5	[redacted]									
Progression Free [redacted]										
1	[redacted]									



EQ-5D-5L item scores range from 1 = No problems, to 5 = Extreme problems/unable to. *Observations.

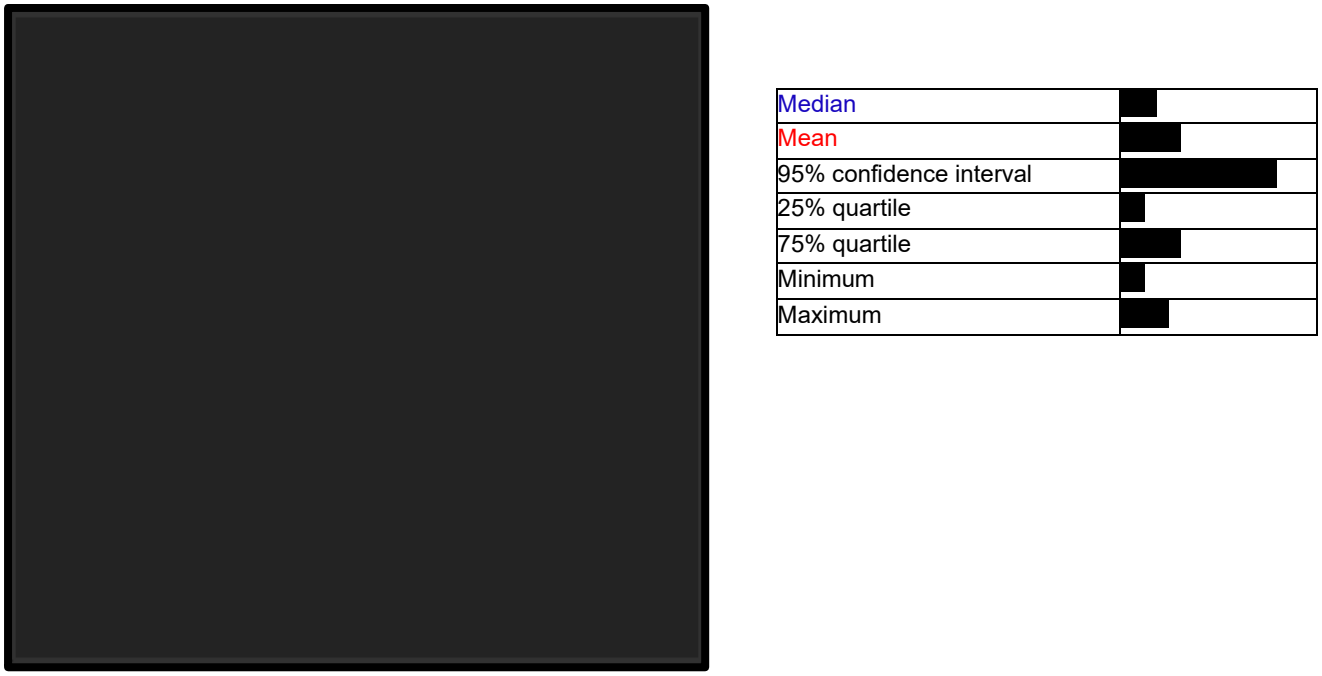
Response distributions between PFS and PD were not vastly different. Therefore, progression does not appear to have had a large impact on HRQoL in the data available. The limited impact on utility associated with progression does not appear to be consistent, given the difference between PFS and PD states observed in other models and appraisals in advanced thyroid cancer, this inconsistency was also validated by UK clinicians in a recent advisory board.[13,26–31]

For example, health state utility values from the DECISION trial of sorafenib in a first-line setting (measured using the EQ-5D-3L) used in multiple-technology appraisal (MTA) (TA535) by the assessment group were 0.72 and 0.80 for patients in PFS receiving sorafenib and BSC, respectively.[32] While individuals in PD state had a utility of 0.64. This equates to a much larger impact associated with progression than that observed in the utility analyses of the COSMIC-311 data. Also, a vignette study by Fordham et al. 2015[33], which aimed to estimate health state utilities in individuals with RR-DTC has also been used and accepted in several NICE appraisals in this clinical area, including TA742[34] in a second-line setting and TA516[35]. In this study, utilities of 0.87 and 0.52 were estimated for the PFS and PD states, respectively.

The limited impact of progression in the COSMIC-311 data was likely a result of limited follow-up in the PD state or missing data, as the data suggests that utility falls over time in the PD state. Regarding missing data, the CSR states that [redacted] progression events occurred before the data cut-off, however only [redacted] participants are captured in the HRQoL assessment after progression. This is due to the fact that HRQoL assessment were discontinued in patients who progressed in the placebo arm and began crossover cabozantinib treatment ([redacted]). Among those captured within the PD HRQoL data, if those in worse health are more likely to drop-out of HRQoL assessments while in the PD state, this could overestimate the progressed disease

value. Figure 26 shows the time between progression and HRQoL assessments in the PD state. The median number of days between progression and HRQoL assessment was [redacted] days, with a mean of [redacted] days. However, the histogram in Figure 26 shows that a large number of PD observations were captured within 10 days of progression ([redacted] within the first 10 days and [redacted] within the first 5 days). If the impact of progression on HRQoL is not immediately felt and increases over time, it is unlikely that the PD utility values obtained from this data will be reflective of the full PD state. Due to this lack of validity of the COSMIC-311 HRQoL data, Fordham et al. 2015[33] utilities were used in the base case.

Figure 26: Histogram showing time from progression to assessment for PD EQ-5D observations (obs)



Note: Blue line denotes median number of days between progression and HRQoL assessment. Red line denotes mean number of days between progression and HRQoL.

To ensure there is no heterogeneity in the population characteristics of COSMIC-311 and Fordham et al. 2015, which may ultimately influence the health state utility values produced, Fordham et al. 2015 utility values have been used for both health states so that utility values come from a consistent data source which share the same population. In addition to this, as per B19, utility values have now been adjusted so that they do not exceed those of the general population. However, the company have provided a scenario whereby PFS utility is sourced from COSMIC-311 and PD is

sourced from Fordham et al. 2015. The results of this scenario are located in the attached Appendix.

B21. PRIORITY. CS, Section B.3.4.3, page 118. Please justify the assumption that all negative HRQoL impacts resulting from AEs are resolved within 1 month.

Response

Negative HRQoL impacts resulting from AEs were applied as a one-off in the first cycle (month) of the model. The decrement applied was based on the mean duration (in days) of that AE, which was obtained from the COSMIC-311 trial data and is shown in Table 11.

Table 11: Mean Duration of AEs, in days

AE	Duration (days)		Reference
	Mean	SD	
Hand-foot syndrome	[redacted]	[redacted]	COSMIC-311
Proteinuria			
Hypertension			
Diarrhoea			
Fatigue			
Hypocalcaemia			

Abbreviations: SD – Standard deviation

The average duration of an AE is [redacted] days. Given this, Ipsen have provided a scenario where negative HRQoL impacts (disutilities) have been applied over two cycles (two months). This update to formula is found in cell AH10 in both the "Trace (Cabozantinib)" and "Trace (BSC)" worksheets. The updated model and appendix with results will be shared on the 26th October.

B22. CS Section B.3.4.3, Table 31. Please clarify the values and source of the disutilities associated with AEs for hand-foot syndrome and diarrhoea, given that these do not match the disutilities used by the ERG in TA535.

Response

We acknowledge the referencing error within the model and can confirm a correction has been made. The correct reference for AEs for hand-foot syndrome and diarrhoea is Fordham et al. (2015)[33] reporting the corresponding values of -0.34 and -0.47, respectively.

Costs

B23. Model, worksheet “Trace (Cabozantinib)”, column I. The model does not include any costs associated with wastage - instead, the model assumes that every tablet prescribed is taken. Please justify the exclusion of wastage costs from the economic model.

Response

As per the response in B24, the base case has been updated to incorporate the EAGs preferred methodology of including wastage (plus the use of RDI over compliance) as per NICE TA474.[36] Wastage is applied in the model assuming 7 days worth of treatment (a quarter of a pack of tablets).

The switch for wastage is located in “Cost Inputs” D23. In cell D24, the user editable wastage cost can be found, for which we have assumed a quarter pack of tablets [redacted]). This cost is then fed through to the “Trace (Cabozantinib)” worksheet in cells I9:I608.

The revised base case results have been provided as part of the Appendix.

B24. PRIORITY. CS, Section 3.5.1.1, page 125. The model calculates acquisition costs as a function of TTD, PAS price and compliance. RDI is not used. Please provide protocol definitions of compliance and RDI in COSMIC-311 and explain why compliance has been used in preference to RDI.

Response

Compliance was calculated as*:*

$$= (days\ of\ follow\ up\ -\ total\ time\ off\ treatment)/\ days\ of\ follow\ up$$

Relative dose intensity (RDI) was calculated as:

$$= 100 * (average\ daily\ dose\ mg/day\ )/(60mg/day)$$

The base case has been updated such that acquisition costs now use the EAGs preferred methodology of RDI with the inclusion of wastage costs, over compliance. Using RDI in the economic model ensures that accurate dosing and number of days

adhered to are captured in the acquisition cost calculation, as opposed to number of days adhered to alone.

A switch ("Settings" G41) has been added in the model which allows the option to model RDI or compliance. This is linked to the "Trace (Cabozantinib)" worksheet in cells I9:I608.

The revised base case results have been provided as part of the Appendix.

B25. CS, Section B.3.5.2.1, page 126. Please justify the assumption that, except for ECGs, cabozantinib will not require any additional monitoring costs over BSC.

Response

The types of resource and frequency of use in the progression free (PF) and progressed disease (PD) health states were based on NICE TA516 and TA742. This includes blood tests that are required as part of monitoring patients with DTC irrespective of whether they are on BSC or not as the disease itself can cause derangements in blood levels of minerals such as calcium which if too low or too high can predispose patients to cardiac arrhythmias. We have actually taken a slightly conservative approach in increasing the frequency of ECG monitoring for cabozantinib in stating monitoring would be performed every two months. The cabozantinib Summary of Product Characteristics (SPC) states an ECG should be done at the beginning of treatment and periodically thereafter, although the actual frequency is not defined in the SPC which could mean ECGs could be done every 6 months rather than every 2 months. Other protein kinase inhibitors used to treat DTC such as lenvatinib and selpercatinib have similar wording in their SPCs. Experts consulted by Ipsen as to how frequently they performed ECGs stated it was done at the start of treatment and then only if clinically indicated.[7] Finally experts remarked that the adverse events for cabozantinib were in line with those expected for protein kinase inhibitors and that the treatment discontinuation rate of 8.8% for cabozantinib was noted to be lower than that reported in the lenvatinib (SELECT) trials which was 14.4%.[7] Lenvatinib was part of TA535 which allocated the same levels of resource use for lenvatinib and sorafenib in the PF and PD health states as BSC.

QALY shortfall

B26. CS, Section B.3.6, pages 130-131. Please clarify which tool was used to calculate the absolute and proportional QALY shortfall.

Response

The CS uses the QALY shortfall calculator published by Schneider et al. (2021)[37]

B27. PRIORITY. The QALY weighting for severity has been applied to the willingness to pay threshold, rather than to the QALY gain. Please provide updated ICERs with the modifier applied directly to the QALYs.

Response

This has been actioned in the company model ("Results", cell I11) with the corresponding ICER presented in "Results", cell K11. The updated model and appendix with results will be shared on the 26th October.

Section C: Textual clarification and additional points

C1. CS, Section B.2.3.1.1, page 23. The text states that 177 patients (rather than 170 patients) received cabozantinib. Please clarify if this is a typographical error.

Response

Yes, this is a typographical error and should be 170 patients.

C2. CS, Section B.2.6.1, Table 11. Please clarify if the duplicate rows for 'Hazard ratio (95% CI; unstratified)[d]' reflect typographical errors.

Response

Through checking the duplicate rows for HR in Table 11 of the CS, we have also found other data discrepancies for PFS and additionally OS in Table 13 of the CS for CCO1. The below tables have been updated accordingly, using the XL184-311 CSR for CCO1, report date 30th April 2021 (19th August 2020 cut-off).

Table 12: Progression-free survival per BIRC (ITT population)

	CCO2* (N = 258)		CCO1** (N = 187)	
	Cabozantinib (N = 170)	Placebo (N = 88)	Cabozantinib (N = 125)	Placebo (N = 62)
Number (%) of subjects				
Censored	[redacted]	[redacted]	[redacted]	[redacted]
Receipt of local radiation to soft tissue for DTC	[redacted]	[redacted]	[redacted]	[redacted]
No post-baseline ATA[a]	[redacted]	[redacted]	[redacted]	[redacted]
No event by last ATA	[redacted]	[redacted]	[redacted]	[redacted]
2 or more missed ATA prior to event	[redacted]	[redacted]	[redacted]	[redacted]
Systemic NPACT	[redacted]	[redacted]	[redacted]	[redacted]
Event	62 (36)	69 (78)	31 (25)	43 (69)
Death	[redacted]	[redacted]	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]	[redacted]	[redacted]
Duration of PFS (months)				
Median (96% CI)	11.0 (7.4, 13.8)	1.9 (1.9, 3.7)	NE (5.7, NE)	1.9 (1.8, 3.6)
25th percentile, 75th percentile[b]	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Observed p-value (stratified log-rank test)[c]	[redacted]		[redacted]	
Hazard ratio (95% CI; stratified)[c,d]	[redacted]		[redacted]	
Hazard ratio (96% CI; stratified)[c,d]	0.22 (0.15, 0.32)		0.22 (0.13, 0.36)	
Observed p-value (unstratified log-rank test)	[redacted]		[redacted]	
Hazard ratio (95% CI; unstratified)[d]	[redacted]		[redacted]	
Hazard ratio (96% CI; unstratified)[d]	[redacted]		[redacted]	
KM landmark estimates (% of subjects event-free) at:				
3 months	[redacted]	[redacted]	[redacted]	[redacted]
6 months	[redacted]	[redacted]	56.9	16.9
9 months	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]

* 8th February 2021 cut-off
** 19th August 2020 cut-off
Abbreviations: ATA – Adequate tumor assessment; BIRC – Blinded independent radiology committee; CI – Confidence interval; DTC – differentiated thyroid cancer; HR – Hazard ratio; ITT – Intent-to-treat; IxRS – Interactive voice/web response system; KM – Kaplan-Meier; NPACT – Nonprotocol anticancer therapy; ORR – Objective response rate; PD – Disease progression; PFS – Progression-free survival
+ indicates a censored observation (please see PFS censoring rules in XL184-311 CSR, Section 9.7.1.2.2)
a. In the Full ITT population, 11 cabozantinib and 8 placebo subjects were enrolled too close to the data cut cutoff date to have had a post-baseline tumour assessment. Four cabozantinib subjects decided to withdraw from

treatment before any postbaseline tumor assessment. In addition, 3 subjects in the cabozantinib arm (1807-3002, 3808-3111, and 3907-3338) and 1 subject in the placebo arm (3905-3275) died before their first post-baseline scan.
b. Percentiles were based on KM estimates.
c. Stratification factors (per IxRS) comprise receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years).
d. Estimated using the Cox proportional-hazard model (adjusted for stratification factors if applicable). HR < 1 indicated PFS in favor of cabozantinib.
Source: XL184-311 CSR (30th April 2020)[23] and XL184-311 CSR Addendum 1 (21st May 2021)[38] and Brose et al, 2021[39]

Table 13: Overall survival CCO1 and CCO2 ITT population

	CCO2*		CCO1**	
	Cabozantinib (N=170)	**Placebo (N=88)**	**Cabozantinib (N=125)**	**Placebo (N=62)**
Number of subjects (%)				
Censored	[redacted]	[redacted]	[redacted]	[redacted]
Alive	131 (77)	67 (76)	NR	NR
Death after data cut-off date	[redacted]	[redacted]	NR	NR
Death	37 (22)	21 (24)	17 (14)	14 (23)
Duration of overall survival (months)[a]				
Median (95% CI)	19.4 (15.9, NE)	NE (NE, NE)	NE (NE, NE)	NE (NE, NE)
25th percentile	[redacted]	[redacted]	[redacted]	[redacted]
75th percentile	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Observed p-value (stratified log-rank test)[b]	[redacted]		[redacted]	
Hazard ratio (95% CI; stratified)[b,c]	0.76 (0.45, 1.31)		0.54 (0.27, 1.11)	
Observed p-value (unstratified log-rank test)	[redacted]		[redacted]	
Hazard ratio (95% CI; unstratified)[c]	[redacted]		[redacted]	
KM landmark estimates (% of subjects event-free) at:				
3 months	[redacted]	[redacted]	[redacted]	[redacted]
6 months	[redacted]	[redacted]	84.8	73.4
9 months	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]

* 8th February 2021 cut-off
** 19th August 2020 cut-off
Abbreviations: CC01- Clinical cut-off 1; CC02-Clinical cut-off 2; CI – Confidence; HR – Hazard ratio, ITT – Intent-to-treat; LR – Log-rank test, NE – Not estimable; NR – Not reported; OS – Overall survival.
+indicates a censored observation (please see OS censoring rules in XL184-311 CSR, Section 9.7.1.4.1).
a Percentiles were based on K-M estimates.
b Stratification factors based on IxRS were receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years).
c Estimated using the Cox proportional-hazard model (adjusted for stratification factors if applicable). HR < 1 indicated OS in
favour of cabozantinib.
d In the Full ITT population and Primary Analysis subset, maximum duration of OS in the placebo arm was 17.28 months at
Source: XL184-311 CSR (30th April 2020)[23] and XL184-311 CSR Addendum 1 (21st May 2021)[38] and Brose et al, 2021[39]

C3. CS, Appendix D.1.2, Table 7. Please clarify if the study 1575 Robinson should correspond to reference 38 instead of reference 48.

Response

Yes, this is correct, the study 1575 Robinson is relating to reference 38 instead of 48.

C4. CS, Appendix F, Table 36. Please clarify the correct data for Grade 3 anorexia (error) and weight loss (missing).

Response

The table below contains the rectified data from the Cabanillas 2017 study.[14]

Table 14: [redacted]

Adverse Event	Grade 1 N (%)	Grade 2 N (%)	Grade 3 N (%)
Clinical (>10% frequency)			
Constitutional			
• Fatigue	[redacted]	[redacted]	[redacted]
• Anorexia			
• Weight loss			
Gastrointestinal (GI)			
• Dysgeusia	[redacted]	[redacted]	[redacted]
• Oral mucositis			
• Dry mouth			
• Nausea			
• Vomiting			
• Diarrhoea			
• Other GI			
Dermatological	[redacted]	[redacted]	[redacted]
• Palmar-plantar erythrodysesthesia	[redacted]	[redacted]	[redacted]
• Rash			
• Other dermatologic disorder			
Vascular			
• Hypertension	[redacted]	[redacted]	[redacted]
• Proteinuria			
• Bleeding			
Other			
• Pain	[redacted]	[redacted]	[redacted]
• Headache			
• Other musculoskeletal			
• Voice alteration			
• Peripheral neuropathy			
Laboratory (all)			

• Liver transaminase elevation • Hypomagnesemia • Lipase or amylase elevation • Hypocalcaemia • Hypophosphatemia • Hyponatremia • Hypokalaemia • Alkaline phosphatase elevation • Hypoalbuminemia • Hyperglycaemia	[redacted]	[redacted]	[redacted]
Hematologic (all) • Anaemia • Thrombocytopenia • Leukopenia • Neutropenia	[redacted]	[redacted]	[redacted]

C5. CS, page 125. Section B.3.5.1.2 states that *"Administration costs were based on NHS References costs 2021/22 83 and PSSRU 2021"*. Please clarify if this is a typographical error, and if the NHS reference cost used was from 2020/21.

Response

We acknowledge the typographical error and can confirm the NHS reference cost used was from 2020/21.[40]

References

1. Durante et al. Cabozantinib versus placebo in patients with radioiodine-refractory differentiated thyroid cancer (DTC) who have progressed after prior VEGFR-Targeted therapy: updated results from the phase 3 COSMIC-311 trial and prespecified subgroup analyses based on prior therapy. Endocrine Abstracts. 2022;
2. Hernando J, Capdevila J, Robinson B, Sherman SI, Jarząb B, Lin CC, et al. Cabozantinib (C) versus placebo (P) in patients (pts) with radioiodine-refractory (RAIR) differentiated thyroid cancer (DTC) who have progressed after prior VEGFR-targeted therapy: Outcomes in prespecified subgroups based on prior VEGFR-targeted therapy. JCO [Internet]. 2022 Jun [cited 2022 Sep 9];40(16_suppl). Available from: https://ascopubs.org/doi/abs/10.1200/JCO.2022.40.16_suppl.6083
3. Capdevila J, Robinson B, Sherman SI, Jarząb B, Lin CC, Vaisman F, et al. Cabozantinib versus placebo in patients (pts) with radioiodine-refractory (RAIR) differentiated thyroid cancer (DTC) who progressed after prior VEGFR-targeted therapy: Outcomes in prespecified subgroups based on histology subtypes. JCO. 2022 Jun;40(16_suppl):6081–6081.
4. Durante C, Robinson B, Sherman SI, Krajewska J, Lin CC, Vaisman F, et al. Effect of age on efficacy and safety of cabozantinib vs placebo in patients with radioiodine refractory (RAI-R)-differentiated thyroid cancer (DTC) with progression after VEGFR-targeted therapy: subgroup analysis from Phase 3 COSMIC 311 study. In: Endocrine Abstracts [Internet]. Bioscientifica; 2022 [cited 2022 Sep 9]. Available from: https://www.endocrine-abstracts.org/ea/0081/ea0081oc3.4

5. Centre for Reviews and Dissemination. Systematic Reviews [Internet [Internet]. 2009; Available from: https://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/SysRev3.htm
6. NICE. Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [ID1059] [Internet]. 2018. Available from: https://www.nice.org.uk/guidance/ta535/chapter/1-Recommendations.
7. Ipsen. IPSEN HTA Advisory Board for Differentiated Thyroid Cancer Insights Report. 2022.
8. Schlumberger M, Tahara M, Wirth LJ, Robinson B, Brose MS, Elisei R. Lenvatinib versus Placebo in Radioiodine-Refractory Thyroid Cancer. N Engl J Med. 2015;Feb 12;372(7):621–30.
9. Brose MS, Nutting CM, Jarzab B, Elisei R, Siena S, Bastholt L. Sorafenib in locally advanced or metastatic, radioactive iodine-refractory, differentiated thyroid cancer: a randomized, double-blind, phase 3 trial. Lancet. 2014;Jul 26;384(9940):319–28.
10. Cancer Research UK: Thyroid cancer incidence statistics. Available at: Thyroid cancer incidence statistics | Cancer Research UK Accessed on: 12th October 2022.
11. Oh HS, Shin DY, Kim M, Park SY, Kim TH, Kim BH. Extended Real-World Observation of Patients Treated with Sorafenib for Radioactive Iodine-Refractory Differentiated Thyroid Carcinoma and Impact of Lenvatinib Salvage Treatment: A Korean Multicenter Study. Thyroid Off J Am Thyroid Assoc. 2019;Dec;29(12):1804–10.
12. Brilli L, Dalmiglio C, Pilli T, Barbato F, Maino F, Capezzone M. Improvement of Overall Survival Using TKIs as Salvage Therapy in Advanced Thyroid

Carcinoma: Real-Life Data on a Single Center Experience. J Clin Med. 2021;Jan 20;10(3):384.

13. Ipsen. IPSEN HTA Advisory Board for Differentiated Thyroid Cancer. [Data on file]. 2022.

14. Cabanillas MEDS, Geyer JA, Wirth S, Menefee LJ, Liu ME, Shah SV. Cabozantinib as salvage therapy for patients with tyrosine kinase inhibitor-refractory differentiated thyroid cancer: Results of a multicenter phase II international thyroid oncology group trial. J Clin Oncol. 2017;35:3315–21.

15. committee-papers.pdf [Internet]. [cited 2022 Oct 25]. Available from: https://www.nice.org.uk/guidance/ta535/documents/committee-papers

16. Schlumberger M. Lenvatinib versus placebo in radioiodine-refractory thyroid cancer. New England Journal of Medicine. 2015;372:621–30.

17. Brose MS, Nutting CM, Jarzab B, Elisei R, Siena S, Bastholt L, et al. Sorafenib in radioactive iodine-refractory, locally advanced or metastatic differentiated thyroid cancer: a randomised, double-blind, phase 3 trial. Lancet. 2014 Jul 26;384(9940):319–28.

18. Latimer NR, Abrams KR. NICE DSU Technical Support Document 16: Adjusting Survival Time Estimates In The Presence Of Treatment Switching [Internet]. School of Health And Related Research; 2014. Available from: https://eclass.uoa.gr/modules/document/file.php/MED1174/Lectures/Unit%207%20%28Missing%20data%20and%20non-compliance%29/TSD16_Treatment_Switching.pdf

19. Howe CJ, Cole SR, Chmiel JS, Muñoz A. Limitation of Inverse Probability-of-Censoring Weights in Estimating Survival in the Presence of Strong Selection Bias. Am J Epidemiol. 2011 Mar 1;173(5):569–77.

20. question-answer-adjustment-cross-over-estimating-effects-oncology-trials_en.pdf [Internet]. [cited 2022 Oct 25]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/question-answer-adjustment-cross-over-estimating-effects-oncology-trials_en.pdf
21. National life tables: England - Office for National Statistics [Internet]. [cited 2022 Oct 11]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables/current
22. Hernández-Alava, Pudney, Wailoo. Estimating EQ-5D by Age and Sex for the UK [Internet]. 2022. Available from: https://nicedsu.sites.sheffield.ac.uk/methods-development/estimating-eq-5d-by-age-and-sex-for-the-uk
23. Exelixis, Inc. Clinical Study Report XL184-311. 30 April 2020. [Data on file].
24. Hernández-Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. Journal of Health Economics. 2017;55(C):139–52.
25. NICE. NICE health technology evaluations: the manual [Internet]. 2022. Available from: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741
26. 2022 ICD-10-CM Diagnosis Code C73: Malignant neoplasm of thyroid gland [Internet]. [cited 2021 Nov 19]. Available from: https://www.icd10data.com/ICD10CM/Codes/C00-D49/C73-C75/C73-/C73
27. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American

Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid. 2016 Jan;26(1):1–133.

28. Fugazzola L, Elisei R, Fuhrer D, Jarzab B, Leboulleux S, Newbold K, et al. 2019 European Thyroid Association Guidelines for the Treatment and Follow-Up of Advanced Radioiodine-Refractory Thyroid Cancer. Eur Thyroid J. 2019 Oct;8(5):227–45.

29. Schmidt A, Iglesias L, Klain M, Pitoia F, Schlumberger MJ. Radioactive iodine-refractory differentiated thyroid cancer: an uncommon but challenging situation. Arch Endocrinol Metab. 2017 Feb;61(1):81–9.

30. Kaae AC, Kreissl MC, Krüger M, Infanger M, Grimm D, Wehland M. Kinase-Inhibitors in Iodine-Refractory Differentiated Thyroid Cancer—Focus on Occurrence, Mechanisms, and Management of Treatment-Related Hypertension. Int J Mol Sci. 2021 Nov 12;22(22):12217.

31. Perros P, Boelaert K, Colley S, Evans C, Evans RM, Gerrard Ba G, et al. Guidelines for the management of thyroid cancer. Clin Endocrinol (Oxf). 2014 Jul;81 Suppl 1:1–122.

32. NICE. Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [ID1059] [Internet]. 2018. Available from: https://www.nice.org.uk/guidance/ta535/documents/committee-papers

33. Fordham BA, Kerr C, de Freitas HM, Lloyd AJ, Johnston K, Pelletier CL, et al. Health state utility valuation in radioactive iodine-refractory differentiated thyroid cancer. Patient Prefer Adherence. 2015;9:1561–72.

34. NICE. Selpercatinib for treating advanced thyroid cancer with RET alterations [ID3744] [Internet]. 2021. Available from: https://www.nice.org.uk/guidance/ta742/documents/committee-papers

35. NICE. Cabozantinib and vandetanib for treating unresectable locally advanced or metastatic medullary thyroid cancer [ID56] [Internet]. 2017. Available from: https://www.nice.org.uk/guidance/ta516/documents/committee-papers

36. NICE. CDF Rapid Reconsideration Sorafenib for advanced hepatocellular carcinoma (review of TA189) [ID1012] Committee Papers [Internet]. 2017. Available from: https://www.nice.org.uk/guidance/ta474/documents/committee-papers-2

37. Schneider P, McNamara, Love-Koh J, Doran T, Gutacker N. QALY Shortfall Calculator. 2021.

38. Exelixis Inc. Clinical Study Report XL184-311 Addendum 1. 21 May 2021. [Data on file].

39. Brose MS, Robinson B, Sherman SI, Krajewska J, Lin CC, Vaisman F, et al. Cabozantinib for radioiodine-refractory differentiated thyroid cancer (COSMIC-311): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Oncology. 2021 Aug 1;22(8):1126–38.

40. NHS England. National Cost Collection: National Schedule of NHS costs - Year 2020-21 - NHS trust and NHS foundation trusts [Internet]. 2022. Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#ncc1819

Ipsen response to follow-up queries from the EAG (received 8th November, 2022)

Question A15

The company's clarification response states that the TTD analysis had missing data and two alternative approaches were applied to estimate TTD. Figure 4 in the response shows TTD estimated using these two alternative approaches. Our questions are:

- Isn't the missing data just censoring?

Response

The patients have not been censored in the truest form as we still have their data for other variables, such as date of progression, despite not always observing their time on treatment. So, we believe this is a missing data issue, and not just censoring.

- Is there something exceptional about this case which wouldn't apply when estimating TTD in other trial datasets?

Response

Time to event data for TTD (discontinuation moment and censoring parameter) were not directly available, as a variable, from the COSMIC-311 data. Therefore, it had to be derived. Variables directly available from COSMIC-311 dataset (ADaM format), "TR01EDT" and "TR01SDT", which are respectively the dates of last and first exposure to treatment in period 01, were used to estimate the TTD for patients for whom both dates were available.

For those with TR01EDT missing, we set their TTD and censoring parameter equal to that of PFS, resulting in the blue curve in the graph. The assumption underlying this curve is that even if the TR01EDT date is not known, patients who progressed will discontinue treatment and those with censored progression get censored discontinuation date as well.

Alternatively, subjects with missing TR01EDT had their TTD and censoring parameter set using the date of last dose ongoing (LTRTOGDT) which is equal to the cut-off date for all patients with missing TR01EDT. The result was the green curve in the graph. This approach accounts for the fact that patients are still under treatment at the cut-off date if the end of treatment date is missing – this was the approach preferred by Ipsen biostats and the one used in the model.

In summary,

- the blue curve was obtained by linking TTD to PFS for subjects who did not have an end of treatment date

- the green curve was obtained by linking TTD to the last observed treatment dosage for subjects who did not have an end of treatment date

- We understand that the green line reflects the data used in the model analysis, and that this links TTD to progression. Under this approach, are events defined as (a) known discontinuation date or (b) known date of progression if discontinuation date is not known, with all other patients censored at their last known dose date (i.e., those with (a) or (b))?

Response

The blue line (not the green) links TTD to progression and the green line relates TTD to the last observed dosage. The curve using last observed dosage was used in the model and events were defined as follows:

1. For subjects with TR01EDT values: time to treatment discontinuation value is equal to TR01EDT – TR01SDT and discontinuation event set to 1
2. For those with TR01EDT missing: time to treatment discontinuation value is equal to LTRTOGDT – TR01SDT and discontinuation event set to 0

The definition is analogous when using the date of progression instead of the last known dose date.

Question B5

Figure 22 in the response shows observed and modelled TTD.

- Why does the Kaplan-Meier estimator in Figure 22 look different to both the green and blue TTD lines in Figure 4?

- We are surprised that all of the models provide such a poor fit. Can the company confirm that the survival models and the KM estimator in Figure 22 relate to the same dataset?

Response

Ipsen recognise that Figure 4 was incorrect as it was data from CCO1 rather than CCO2 for cabozantinib.

Please see below updated figure (Figure 1) for the CCO2 TTD data.

Blue line = TTD using date of progression
Green line = TTD using last observed dose

Figure 1: Updated KM curves for PFS and TTD for CCO2

Ipsen apologises for the confusion, the TTD coefficients used in the model are using last observed dose from CCO2 for cabozantinib, whereas the KM data was TTD linked with the date of progression. Figure 2 now reflects the updated KM curves figure (green line in Figure 1 above). The TDD curves now show a good fit with the KM data.

Figure 2: TTD curves from the model for CCO2 for cabozantinib

APPENDIX: EAG Clarification Questions – Ipsen response (26/10/22)

Base-case results

Base-case incremental cost-effectiveness analysis results

The updates to the revised base case include:

- Corrections as per B17 (Half-cycle correction, updated life tables, general population mortality risk constraint, AE frequencies and cap on utility values)
- Replacing compliance with RDI plus wastage as per B24
- Updating the data source used for BSC OS to be CCO1 as per B6
- Removal of 5 year BSC OS constraint as per B6

As described in the Company Submission (CS), a confidential simple patient access scheme (PAS) has been approved by the Patient Access Schemes Liaison Unit (PASLU). The pack price under this scheme is [redacted] (a [redacted]). This PAS has been applied and the results presented to reflect this discount. As per Section B.3.6 in the CS, locally advanced or metastatic DTC patients, refractory or not eligible to RAI who have progressed during or after prior systemic therapy qualifies for the 1.2 severity modifier. The modifier of 1.2 has been applied to the incremental QALYs. The deterministic, base case incremental cost-effectiveness analysis results are presented in Table 1. Cabozantinib was associated with [redacted] incremental costs and [redacted] incremental QALYs compared to BSC, which corresponds to an ICER of £20,289 per QALY gained. Disaggregated base case results are presented in Table 3, Table 4 and Table 5.

The net health benefit (NHB) is displayed in Table 2. The NHB at £30,000 of 0.361 implies that overall population health would be increased as a result of introducing cabozantinib.

Table 1: Deterministic base-case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	ICER versus baseline (£/QALY)
PAS price							
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,289

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Table 2: Net health benefit

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs*	NHB at £20,000	NHB at £30,000
PAS price						
BSC	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	-0.016	0.361

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; NHB – Net health benefit; QALYs – Quality-adjusted life years

A summary of QALY gain by health state is presented in Table 3. The largest QALY increment between cabozantinib and BSC was observed in the PFS health state.

Table 3: Summary of QALY gain by health state

Health state	QALY Cabozantinib	QALY BSC	Increment Cabozantinib vs. BSC*	Absolute increment Cabozantinib vs. BSC	% absolute increment Cabozantinib vs. BSC
PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; PFS – Progression free survival; PD – Progressed disease; QALY – Quality adjusted life year.

A summary of the costs by health state is presented in Table 4. The largest increment between cabozantinib and BSC was observed in the PFS health state.

Table 4: Summary of costs by health state

Health state	Cost Cabozantinib	Cost BSC	Increment Cabozantinib vs. BSC	Absolute increment Cabozantinib vs. BSC	% absolute increment Cabozantinib vs. BSC
PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dead	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC – Best supportive care; PFS – Progression free survival; PD – Progressed disease.

A summary of the predicted resource use by category of cost is presented in Table 5. The largest increment between cabozantinib and BSC was due to the treatment costs.

Table 5: Summary of predicted resource use by category of cost

Item	Cost Cabozantinib	Cost BSC	Increment Cabozantinib vs. BSC	Absolute increment Cabozantinib vs. BSC	% absolute increment Cabozantinib vs. BSC
Treatment cost (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Health state cost (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event cost (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total cost (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC – Best supportive care.

Exploring uncertainty

Probabilistic sensitivity analysis

The mean values for total costs, LYs, QALYs, and incremental cost per QALY gained for cabozantinib versus BSC for the population of interest generated through 10,000 simulations of the base-case PSA are presented in Table 6. The output shows that on average, cabozantinib results in [redacted] incremental QALYs compared to BSC. In addition, cabozantinib is associated with [redacted] incremental costs over a life-time horizon compared with BSC, resulting in an ICER of £20,515.

Figure 1 to Figure 3 display the ICEP, CEAC and CEAF of cabozantinib versus BSC. The probabilistic results are centred around the deterministic value and the CEAC shows that from a willingness to pay threshold of £24,000, cabozantinib is cost-effective.

Table 6: Probabilistic sensitivity analyses – Base case

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-		-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 1: Incremental cost-effectiveness plane– Base case

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 2: Cost-effectiveness acceptability curve – Base case

Abbreviations: BSC – Best supportive care

Figure 3: Cost-effectiveness acceptability frontier – Base case

Abbreviations: BSC – Best supportive care

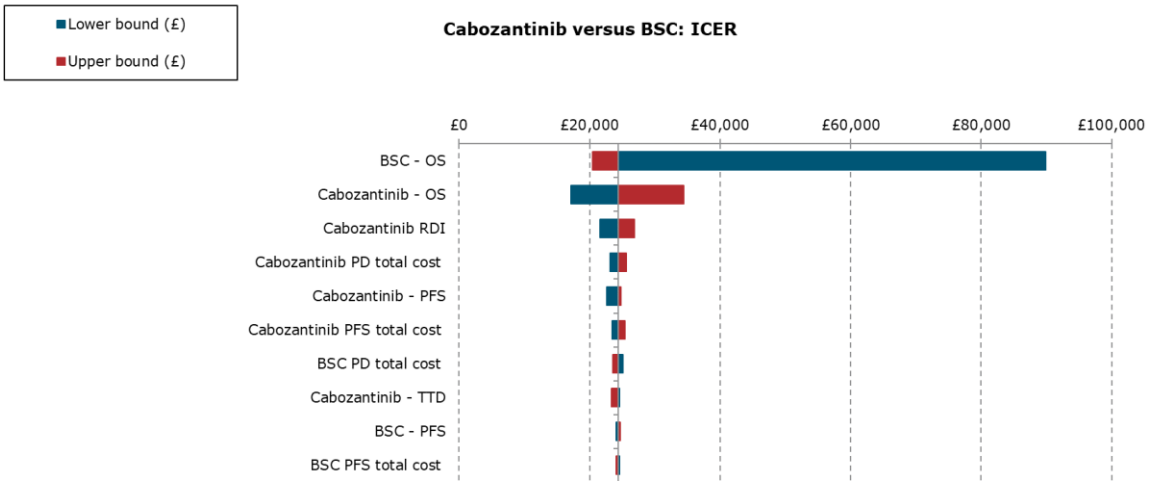
Deterministic sensitivity analysis

Deterministic one-way sensitivity analysis (OWSA) was conducted to explore the level of uncertainty in the model results. The OWSA involved varying one parameter at a time and assessing the subsequent impact on the incremental QALYs and incremental costs. By adjusting each parameter individually, the sensitivity of the model results to that parameter can be assessed.

The OWSA was conducted by allocating a 'low' value and a 'high' value to each parameter; the low value is the lower bound of the 95% CI, the high value is the upper bound of the 95% CI. In the absence of CI data, the variable was altered by +/- 10%. A tornado diagram was developed to graphically present the parameters which have the greatest effect on the ICER.

A OWSA tornado diagram presenting the top 10 most sensitive parameters for cabozantinib versus BSC is presented in Figure 4. Table 7 presents the OSWA results for these 10 parameters. The model was most sensitive to the overall survival of BSC and cabozantinib.

Figure 4: One-way sensitivity analysis tornado plot



Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; OS – Overall survival; PD – Progressed disease; PFS – Progression free survival; TTD – Time to treatment discontinuation

Table 7: One-way sensitivity analysis results

Parameter	Lower bound (£)	Upper bound (£)	Difference (£)
BSC - OS	£89,850	£20,471	£69,378
Cabozantinib - OS	£17,159	£34,437	£17,278
Cabozantinib RDI	£21,584	£26,852	£5,268
Cabozantinib PD total cost	£23,153	£25,663	£2,510
Cabozantinib - PFS	£22,625	£24,761	£2,136
Cabozantinib PFS total cost	£23,422	£25,366	£1,945
BSC PD total cost	£25,060	£23,561	£1,499
Cabozantinib - TTD	£24,630	£23,291	£1,339
BSC - PFS	£24,061	£24,674	£613
BSC PFS total cost	£24,618	£24,049	£569

Abbreviations: BSC – Best supportive care; OS – Overall survival; PD – Progressed disease; PFS – Progression free survival; TTD – Time to treatment discontinuation

Scenario analysis

Table 8 details deterministic scenario analysis results for cabozantinib versus BSC. Cabozantinib is cost-effective at the £30,000 per QALY threshold in all scenarios.

Table 8: Deterministic scenario analysis results

Description	Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Severity modified incremental QALYs	Severity ICER incremental (£/QALY)
Base case	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,289
Discount rate: 0%	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,273
Discount rate: 5%	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,710
Age adjusted utilities: excluded	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,013
PFS: Exponential	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,997
PFS: Generalized gamma	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,402
PFS: Gompertz	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,907
PFS: Log logistic	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,362

Description	Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Severity modified incremental QALYs	Severity ICER incremental (£/QALY)
PFS: Log normal	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,494
OS: Weibull	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	23,669
CCO1 for OS: Both treatment arms	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,299
BSC OS: CCO2 with 5-year OS constraint	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	26,543
Crossover method: Two-stage	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,694
BSC OS: CCO1 with 5-year OS constraint: enabled	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,993
Dosing: Compliance and no wastage	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,954
AE HRQoL impact: 2 cycles	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,445
PFS utility: COSMIC-311	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,382

Probabilistic results

In all scenarios cabozantinib is cost-effective at the £30,000 per QALY threshold.

Discount rate – 0%

Table 9: Probabilistic scenario analysis results - Discount rate 0%

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,458

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 5: Incremental cost-effectiveness plane - Discount 0%

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 6: Cost-effectiveness acceptability curve - Discount 0%

Abbreviations: BSC – Best supportive care

Figure 7: Cost-effectiveness acceptability frontier - Discount 0%

Abbreviations: BSC – Best supportive care

Discount rate – 5%

Table 10: Probabilistic scenario analysis results - Discount rate 5%

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,050

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 8: Incremental cost-effectiveness plane - Discount 5%

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 9: Cost-effectiveness acceptability curve - Discount 5%

Abbreviations: BSC – Best supportive care

Figure 10: Cost-effectiveness acceptability frontier - Discount 5%

Abbreviations: BSC – Best supportive care

Age-adjusted utilities – excluded

Table 11: Probabilistic scenario analysis results – Age-adjusted utilities excluded

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,304

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 11: Incremental cost-effectiveness plane - Age-adjusted utilities excluded

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 12: Cost-effectiveness acceptability curve - Age-adjusted utilities excluded

Abbreviations: BSC – Best supportive care

Figure 13: Cost-effectiveness acceptability frontier - Age-adjusted utilities excluded

Abbreviations: BSC – Best supportive care

PFS: Exponential

Table 12: Probabilistic scenario analysis results – PFS exponential

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,070

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS –Progression free survival; QALYs – Quality-adjusted life years

Figure 14: Incremental cost-effectiveness plane – PFS exponential

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 15: Cost-effectiveness acceptability curve – PFS exponential

Abbreviations: BSC – Best supportive care

Figure 16: Cost-effectiveness acceptability frontier - PFS exponential

Abbreviations: BSC – Best supportive care

PFS: Generalized gamma

Table 13: Probabilistic scenario analysis results – PFS generalized gamma

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,959

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS – Progression free survival; QALYs – Quality-adjusted life years

Figure 17: Incremental cost-effectiveness plane – PFS generalized gamma

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 18: Cost-effectiveness acceptability curve – PFS generalized gamma

Abbreviations: BSC – Best supportive care

Figure 19: Cost-effectiveness acceptability frontier – PFS generalized gamma

Abbreviations: BSC – Best supportive care

PFS: Gompertz

Table 14: Probabilistic scenario analysis results – PFS Gompertz

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	11,492

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 20: Incremental cost-effectiveness plane – PFS Gompertz

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 21: Cost-effectiveness acceptability curve – PFS Gompertz

Abbreviations: BSC – Best supportive care

Figure 22: Cost-effectiveness acceptability frontier – PFS Gompertz

Abbreviations: BSC – Best supportive care

PFS: Log logistic

Table 15: Probabilistic scenario analysis results – PFS loglogistic

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,108

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 23: Incremental cost-effectiveness plane – PFS loglogistic

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 24: Cost-effectiveness acceptability curve – PFS loglogistic

Abbreviations: BSC – Best supportive care

Figure 25: Cost-effectiveness acceptability frontier – PFS loglogistic

Abbreviations: BSC – Best supportive care

PFS: Log normal

Table 16: Probabilistic scenario analysis results – PFS log normal

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,176

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS –Progression free survival; QALYs – Quality-adjusted life years

Figure 26: Incremental cost-effectiveness plane – PFS log normal

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 27: Cost-effectiveness acceptability curve – PFS log normal

Abbreviations: BSC – Best supportive care

Figure 28: Cost-effectiveness acceptability frontier – PFS log normal

Abbreviations: BSC – Best supportive care

OS: Weibull

Table 17: Probabilistic scenario analysis results – OS Weibull

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; OS – Overall survival; QALYs – Quality-adjusted life years

Figure 29: Incremental cost-effectiveness plane - OS Weibull

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 30: Cost-effectiveness acceptability curve - OS Weibull

Abbreviations: BSC – Best supportive care

Figure 31: Cost-effectiveness acceptability frontier - OS Weibull

Abbreviations: BSC – Best supportive care

CCO1 for OS: Both treatment arms

Table 18: Probabilistic scenario analysis results – CCO1 both treatment arms

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; OS – Overall survival; QALYs – Quality-adjusted life years

Figure 32: Incremental cost-effectiveness plane – CCO1: both treatment arms

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 33: Cost-effectiveness acceptability curve – CCO1 both treatment arms

Abbreviations: BSC – Best supportive care

Figure 34: Cost-effectiveness acceptability frontier - CCO1 both treatment arms

Abbreviations: BSC – Best supportive care

BSC OS: CCO2 plus 5-year constraint

Table 19: Probabilistic scenario analysis results – BSC OS: CCO2 with 5-year OS constraint

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; OS – Overall survival; QALYs – Quality-adjusted life years

Figure 35: Incremental cost-effectiveness plane – BSC OS: CCO2 with 5-year OS constraint

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 36: Cost-effectiveness acceptability curve – BSC OS: CCO2 with 5-year OS constraint

Abbreviations: BSC – Best supportive care

Figure 37: Cost-effectiveness acceptability frontier – BSC OS: COO2 with 5-year OS constraint

Abbreviations: BSC – Best supportive care

Crossover method: Two-stage

Table 20: Probabilistic scenario analysis results – Crossover method: Two-stage

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,894

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 38: Incremental cost-effectiveness plane – Crossover method: Two-stage

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 39: Cost-effectiveness acceptability curve – Crossover method: Two-stage

Abbreviations: BSC – Best supportive care

Figure 40: Cost-effectiveness acceptability frontier – Crossover method: Two-stage

Abbreviations: BSC – Best supportive care

BSC OS: CCO1 with 5-year OS constraint: enabled

Table 21: Probabilistic scenario analysis results – BSC OS: CCO1 with 5-year OS constraint: enabled

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,916

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 41: Incremental cost-effectiveness plane – BSC OS: CCO1 with 5-year OS constraint: enabled

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 42: Cost-effectiveness acceptability curve – BSC OS: CCO1 with 5-year OS constraint: enabled

Abbreviations: BSC – Best supportive care

Figure 43: Cost-effectiveness acceptability frontier BSC OS: CCO1 with 5-year OS constraint: enabled

Abbreviations: BSC – Best supportive care

Compliance dosing and no wastage

Table 22: Probabilistic scenario analysis results – Compliance dosing and no wastage

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,038

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 44: Incremental cost-effectiveness plane – compliance dosing and no wastage

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 45: Cost-effectiveness acceptability curve – compliance dosing and no wastage

Abbreviations: BSC – Best supportive care

Figure 46: Cost-effectiveness acceptability frontier – compliance dosing and no wastage

Abbreviations: BSC – Best supportive care

AE HRQoL impact: 2 cycles

Table 23: Probabilistic scenario analysis results – AE HRQoL impact: 2 cycles

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,630

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 47: Incremental cost-effectiveness plane – AE HRQoL impact: 2 cycles

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 48: Cost-effectiveness acceptability curve – AE HRQoL impact: 2 cycles

Abbreviations: BSC – Best supportive care

Figure 49: Cost-effectiveness acceptability frontier – AE HRQoL impact: 2 cycles

Abbreviations: BSC – Best supportive care

PFS utility: COSMIC-311

Table 24: Probabilistic scenario analysis results – PFS utility: COSMIC-311

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs*	Cost per QALY
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,882

*Severity modifier of 1.2 has been applied.
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; PFS –Progression free survival; QALYs – Quality-adjusted life years

Figure 50: Incremental cost-effectiveness plane – PFS utility: COSMIC-311

Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 51: Cost-effectiveness acceptability curve – PFS utility: COSMIC-311

Abbreviations: BSC – Best supportive care

Figure 52: Cost-effectiveness acceptability frontier – PFS utility: COSMIC-311

Abbreviations: BSC – Best supportive care

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]
External Assessment Group Report

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK Christopher Carroll, Reader in Systematic Review and Evidence Synthesis, ScHARR, University of Sheffield, Sheffield, UK Kate (Shijie) Ren, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Aline Navega Biz, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Sarah (Sa) Ren, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Mark Clowes, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK
Correspondence Author	Aline Navega Biz, Research Associate, ScHARR, University of Sheffield, Sheffield, UK
Date completed	8th December 2022 (post-FAC version)

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135663.

Declared competing interests of the authors

None of the authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Dr Caroline Brammer (Clatterbridge Cancer Centre NHS Foundation Trust) and Dr Mike Flatley (Leeds Teaching Hospitals NHS Trust) for clinical advice relating to this project. We would also like to thank Kate Ennis, Research Associate, ScHARR, for providing comments on the draft report and Andrea Shippam, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Tappenden P, Carroll C, Ren K, Navega Biz A, Ren S, Clowes M. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. External Assessment Group report. School of Health and Related Research (ScHARR), 2022.

Contributions of authors

Mark Clowes critiqued the company's search strategy. Christopher Carroll summarised and critiqued the clinical effectiveness data reported within the company's submission. Kate Ren and Sarah Ren critiqued the statistical aspects of the submission. Aline Navega Biz and Paul Tappenden critiqued the health economic analysis submitted by the company. All authors were involved in drafting and commenting on the final report.

CONTENTS

List of tables

List of figures

List of boxes

Abbreviations

A&E	Accident and Emergency
AE	Adverse event
AF	Acceleration factor
AIC	Akaike Information Criterion
ALT	Alanine aminotransferase
ASA	Additional sensitivity analysis
AST	Aspartate aminotransferase
BIC	Bayesian Information Criterion
BIRC	Blinded Independent Radiology Committee
BOCR	Best overall confirmed response
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CCO	Clinical cut-off
CDF	Cancer Drugs Fund
CEAC	Cost-effectiveness acceptabiility curve
CI	Confidence interval
CONSORT	Consolidated Standards of Reporting Trials
CR	Complete response
CRD	Centre for Reviews and Dissemination
CS	Company's submission
CSR	Clinical study report
CT	Computerised tomography
DoR	Duration of response
DSA	Determinsitic sensitivity analysis
DSU	Decision Support Unit
DTC	Differentiated thyroid cancer
EA	Exploratory analysis
EAG	External Assessment Group
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EMBASE	Excerpta Medica Database
EQ-5D-5L	Euroqol 5-Dimensions (5-Level)
ESMO	European Society for Medical Oncology
FTC	Follicular thyroid carcinoma
GP	General practitioner
HCHS	Hospital and Community Health Services
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IPCW	Inverse probability censoring weighting
IPD	Individual patient data
IQR	Interquartile range
ITC	Indirect treatment comparison
ITT	Intention-to-treat
MEDLINE	Medical Literature Analysis and Retrieval System Online
Mg	Milligram
MHRA	Medicines and Healthcare products Regulatory Agency
N/a	Not applicable
NHB	Net health benefit
NHS	National Health Service
NHSCII	NHS Cost Inflation Index
NICE	National Institute for Health and Care Excellence

NMA	Network meta-analysis
NTRK	Neurotrophic tyrosine receptor kinase
OITT	Overall response rate intention-to-treat
ONS	Office for National Statistics
ORR	Objective response rate
OS	Overall survival
PAS	Patient Access Scheme
PFS	Progression-free survival
PH	Proportional hazards
PK	Pharmacokinetic
PPES	Palmar-plantar erythrodysaesthesia syndrome
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTC	Papillary thyroid carcinoma
QALY	Quality-adjusted life year
QD	*quaque die* (once daily)
RAI	Radioactive iodine
RCC	Renal cell carcinoma
RCS	Restricted cubic spline
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
RET	Rearranged during transfection
RPSFT	Rank-preserving structural failure time
RTK	Receptor tyrosine kinases
SAE	Serious adverse event
SE	Standard error
SEER	Surveillance, Epidemiology, and End Results
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
TA	Technology Appraisal
TC	Thyroid cancer
TEAE	Treatment-emergent adverse event
TKI	Tyrosine kinase inhibitor
TSD	Technical Support Document
TSE	Two-stage estimation
TSH	Thyroid-stimulating hormone
TTD	Time to treatment discontinuation
TTO	Time trade-off
UK	United Kingdom
VAS	Visual Analogue Scale
VEGFR	Vascular endothelial growth factor receptor
WTP	Willingness-to-pay
YHEC	York Health Economics Consortium

1. EXECUTIVE SUMMARY

This report assesses cabozantinib for the treatment of adult patients with locally advanced or metastatic differentiated thyroid cancer (DTC), whose disease is refractory to, or who are unsuitable for radioactive iodine (RAI), and whose disease has progressed during or after prior systemic therapy. This summary provides a brief overview of the key issues identified by the External Assessment Group (EAG) as being potentially important for decision-making. It also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.5 explain the key issues in more detail. The results of the EAG's preferred analysis are summarised in Section 1.6. Background information on the condition, technology and evidence and information on non-key issues are detailed in the main EAG report.

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

The key issues identified by the EAG are summarised in Table 1.

Table 1: Summary of the EAG's key issues

ID4046	Summary of issue	Report sections
Issue 1	Uncertainty around the effect of cabozantinib on overall survival	5.3.5 (critical appraisal point [4])
Issue 2	Uncertainty around the most appropriate health state utility values	5.3.5 (critical appraisal point [5])
Issue 3	Issues relating to resource use and costs	5.3.5 (critical appraisal point [6])

There are three key differences between the company's original base case analysis and the EAG's preferred analysis:

(i) Overall survival. The company's model uses exponential distributions fitted to data from the COSMIC-311 trial to estimate overall survival (OS) for both treatment groups, including a structural assumption that all patients receiving best supportive care (BSC) who remain alive at 5 years will die at this timepoint. The EAG's preferred analysis removes the 5-year death assumption for the BSC group.

(ii) Health state utility values. The company's base case model uses utility values from a time trade-off (TTO) study of health states in RAI-refractory DTC. The values used by the company are based on an adjusted regression analysis. The EAG's preferred analysis uses the observed mean values from this study.

(iii) Cost assumptions. The company's model implicitly assumes a stopping rule at progression and excludes drug wastage. The EAG's preferred model removes the stopping rule and includes wastage.

1.2 Overview of key model outcomes

NICE technology appraisals (TAs) compare how much a new technology improves length of life and health-related quality of life (HRQoL) in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Compared with BSC alone, cabozantinib is assumed to impact on QALYs by:

- Extending progression-free survival (PFS)
- Extending OS
- Increasing the frequency of adverse events (AEs), which leads to greater QALY losses compared with BSC.

Compared with BSC alone, cabozantinib is assumed to affect costs by:

- Increasing overall costs due to the acquisition cost of cabozantinib
- Increasing overall disease management costs due to extended OS
- Increasing costs associated with managing AEs.

The modelling assumptions that have the greatest effect on the ICER for cabozantinib versus BSC are:

- The approach used to model OS in each treatment group
- The choice of utility values applied to the progression-free and progressed disease health states
- The inclusion of post-progression cabozantinib costs.

1.3 The decision problem: Summary of the EAG's key issues

Current recommendations from NICE for first-line treatment of RAI-refractory DTC involves systemic therapy with a vascular endothelial growth factor receptor (VEGFR) tyrosine kinase inhibitor (TKI) – either sorafenib or lenvatinib. NICE has also issued positive recommendations for selpercatinib, entrectinib and larotrectinib; however, these treatments are only available for DTC patients with specific mutations through the Cancer Drugs Fund (CDF). There is currently no routinely commissioned second-line treatment for patients with RAI-refractory DTC who have progressed during or after prior systemic therapy. The company's proposed positioning for cabozantinib is in line with its licensed indication for the DTC indication, that is, as monotherapy for the treatment of adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed during or after prior systemic therapy. The decision problem addressed in the company's submission (CS) is generally in line with the final NICE scope.

The EAG notes that whilst the NICE scope includes BSC as the only comparator, some clinicians offer continued lenvatinib after disease progression; however, it is unlikely that sufficient evidence exists to undertake a comparison of cabozantinib against continued post-progression lenvatinib.

1.4 The clinical effectiveness evidence: Summary of the EAG's key issues

The clinical evidence presented in the CS was informed by a systematic literature review (SLR) of studies assessing the clinical efficacy and safety of cabozantinib in adult patients with RAI-refractory DTC receiving second- or third-line treatment, who have been previously treated with sorafenib and/or lenvatinib. The primary clinical evidence detailed in the CS comes from COSMIC-311. This was an international Phase III, multi-centre, placebo-controlled, blinded, randomised controlled trial (RCT), which assessed cabozantinib at the licensed dose of 60mg daily. The trial had two clinical cut-offs (CCOs): the primary cut-off date was the 19th of August 2020 (number of patients: 125 cabozantinib, 62 placebo) (CCO1); and, after further enrolment, the second 'supportive analyses' cut-off date was the 8th of February 2021 (170 cabozantinib, 88 placebo) (CCO2).

COSMIC-311 was a medium-sized trial with 258 subjects across two arms at CCO2, but with a short length of follow-up (median 10.1 months at the latest data cut-off, CCO2, and 6.2 months at the primary data cut-off, CCO1). Cabozantinib demonstrated significant efficacy compared with placebo in terms of PFS and objective response rate (ORR) at both data cut-offs. The study was assessed by the EAG as being at high risk of bias on account of the deviation from the pre-specified interventions: sizeable proportions of patients with progressive disease in the placebo arm crossed-over to receive open-label treatment with cabozantinib therapy within a median period of only 1.9 months after commencement of the trial (31% at CCO1 and 45% at CCO2). This potentially confounded the outcomes of OS and safety. The CS accepts that there was no significant difference between the two arms in terms of OS, only a trend favouring cabozantinib, even after adjusting for treatment switching. Meta-analysis was not conducted, despite the existence of a single-arm trial that satisfied the SLR criteria (NCT01811212). Indirect treatment comparisons (ITCs) were not undertaken due to the absence of comparable trials of second-line therapy in the target population, and the availability of direct evidence from COSMIC-311.

There were high rates of treatment-related AEs and serious adverse events (SAEs) in the cabozantinib arm compared with the placebo arm, as well as dose modifications due to AEs. A number of AEs related to cabozantinib treatment were frequent: diarrhoea, palmar-plantar erythrodysaesthesia syndrome (PPES), hypertension, fatigue, hypocalcaemia, [redacted], and decreased appetite. Some of these AEs were also the most frequent events at Grade ≥3. HRQoL was only assessed by the Euroqol 5-Dimensions 5-Level (EQ-5D-5L) questionnaire in patients who had not progressed / up to the point of progression (to prevent confounding due to crossover) and no significant or clinically important difference between cabozantinib and placebo was found for patients who had not progressed up to 33 weeks (there were only five or fewer patients in the placebo arm after this point, preventing meaningful comparisons from being made).

Issues relating to the clinical evidence also impact on the company's economic analysis; hence, all issues are discussed together in Section 1.5.

1.5 The cost-effectiveness evidence: Summary of the EAG's key issues

The company's economic model assesses the cost-effectiveness of cabozantinib (plus BSC) versus BSC alone for adult patients with locally advanced or metastatic DTC, whose disease is refractory to, or who are unsuitable for RAI, and whose disease has progressed during or after prior systemic therapy. The model adopts a partitioned survival approach which includes three health states: (i) progression-free; (ii) progressed disease and (iii) dead. The analysis adopts an NHS and Personal Social Services (PSS) perspective, including QALYs accrued by DTC patients; caregiver effects are not included. Clinical outcomes for both treatment groups are modelled using parametric survival distributions fitted to data on PFS and OS from COSMIC-311 (CCO2), including adjustment of OS to account for treatment switching which occurred in the placebo arm of the trial. The model includes a structural assumption that all patients in the BSC group who remain alive at 5 years will die at this timepoint. Health state utility values are based on estimates reported from an external TTO valuation study of RAI-refractory DTC health states (Fordham *et al.*). Resource use and cost parameters are based on data from COSMIC-311, clinical input obtained within previous NICE technology appraisals (TAs), other literature and standard costing sources.

A Patient Access Scheme (PAS) is available for cabozantinib which takes the form of a simple price discount of [redacted] (PAS price = [redacted] for 30 days' supply). All results presented in this EAG report include this PAS. Excluding QALY weighting, the probabilistic version of the company's model suggests that compared with BSC, cabozantinib generates an additional [redacted] QALYs at an additional cost of [redacted]; the corresponding ICER is £27,169 per QALY gained. The company's QALY shortfall calculations suggest a decision modifier of 1.2. When QALY weighting is included, the probabilistic ICER is estimated to be £22,641 per QALY gained. The ICERs generated using the deterministic version of the model are slightly higher (ICER excluding QALY weighting = £28,148 per QALY gained; ICER including QALY weighting = £23,456 per QALY gained).

As part of their response to clarification questions from the EAG, the company submitted a revised base case model which re-estimates OS for the BSC group using data from the earlier CCO1 data-cut of COSMIC-311, but which retains the CCO2 data-cut for the cabozantinib group. This model includes a number of error corrections and alternative assumptions. The probabilistic version of the company's revised model suggests a lower ICER than their original model (excluding QALY weighting ICER = £24,616 per QALY gained).

The EAG's has three key concerns regarding the company's original model which relate to: the approach used to model OS (Issue 1); the health state utility values (Issue 2) and the resource use and costing assumptions (Issue 3). These issues are summarised below.

Issue 1: Uncertainty around the effect of cabozantinib on overall survival

Report section	5.3.5 (critical appraisal point [4])
Description of issue and why the EAG has identified it as important	The company's economic model estimates OS for cabozantinib and BSC using parametric survival models which have been fitted to data from COSMIC-311. The exponential distribution was selected for use in the company's base case analysis because: (a) most of the other standard parametric models resulted in OS functions for cabozantinib and BSC which cross - this was considered implausible by the company's clinical advisors, and (b) a hybrid Kaplan-Meier-model with an exponential tail was used to estimate OS in NICE Technology Appraisal No 535 (TA535). The model also includes a structural assumption that all BSC-treated patients who are alive at 5 years will die at this timepoint. The EAG considers the company's approach to modelling OS to be problematic for several reasons: (i) The exponential model does not provide a good representation of the observed Kaplan-Meier estimates or the underlying empirical hazards. (ii) The use of exponential distributions in each group implies an indefinite relative treatment effect (a constant hazard ratio [HR]). The empirical time-varying HR for OS in COSMIC-311 crosses 1.0 after around 6 months and subsequently suggests a higher risk of death for cabozantinib-treated patients compared with BSC-treated patients after this timepoint. (iii) The 5-year death assumption for the BSC group leads to a vertical drop in modelled OS which is unrealistic. The EAG believes that this assumption was included because the company's clinical advisors commented that the exponential model overestimates long-term OS. However, excluding this assumption results in OS estimates which are consistent with the EAG's clinical advisors' expectations of OS for BSC. The company's revised base case model uses earlier OS data from CCO1 for BSC and CCO2 for cabozantinib and removes the 5-year death assumption. However, the EAG does not believe that this model is suitable for decision-making as it attempts to improve the plausibility of the OS predictions for BSC by discarding data from the placebo group of COSMIC-311.
What alternative approach has the EAG suggested?	The EAG asked the company to explore the use of more flexible parametric survival distributions (e.g., restricted cubic spline [RCS] models); however, the company did not fit these models as they anticipated that the OS functions would more closely follow the empirical hazards, thereby leading to the OS models crossing. The EAG has undertaken four exploratory analyses around OS: • *EAG6*: Exponential models excluding the 5-year BSC death assumption • *ASA1a*: Same as EA6 but with treatment effect waning assumed at 3 years • *ASA1b*: Hybrid Kaplan-Meier function plus exponential tail after 12 months, including a constant HR • *ASA1c*: Hybrid Kaplan-Meier function plus exponential tail after 12 months, BSC hazard rate in both treatment groups (HR=1.0).
What is the expected effect on the cost-effectiveness estimates?	Including QALY weighting, the company's original deterministic base case ICER is £23,456 per QALY gained. The EAG's preferred analysis, which removes the 5-year death assumption and includes additional amendments, suggests a higher deterministic ICER of £32,397 per QALY gained (EA6a). The EAG's sensitivity analyses around OS also lead to comparatively higher ICERs: • ASA1a (exponential + 3-year effect waning) = £39,989 per QALY gained • ASA1b (KM_{12} + exponential tail, constant HR) = £33,895 per QALY gained • ASA1c (KM_{12} + exponential tail, HR=1.0) = £59,240 per QALY gained.
What additional evidence or analyses might help to resolve this key issue?	The long-term effect of cabozantinib on OS is highly uncertain and this is a key driver of the ICER. None of the economic analyses presented by the company or the EAG are ideal. Longer-term follow-up in COSMIC-311 would help to reduce uncertainty around long-term OS estimates for cabozantinib and BSC. However, there are no further planned data-cuts of the trial beyond CCO2.

Issue 2: Uncertainty around the most appropriate health state utility values

Report section	5.3.5 (critical appraisal point [5])
Description of issue and why the EAG has identified it as important	The company's model applies utility values based on adjusted estimates obtained from a multivariable regression analysis of response data from a TTO exercise reported by Fordham *et al.* (progression-free utility value = 0.87; progressed disease utility value = 0.52). QALY losses are also applied to account for the impact of AEs. The EAG has several concerns regarding the health state utility values applied in the company's base case model: (i) The TTO vignette method used in Fordham *et al.*, is not in line with the NICE Reference Case. (ii) The utility value for the progression-free state is higher than UK general population norms (0.87 versus 0.82). This implies that it is better to have the disease than not have the disease. (iii) Most previous NICE appraisals of treatments for DTC have applied lower utility values from Fordham *et al.* based on the observed mean values (progression-free utility value = 0.80; progressed disease utility value = 0.50). (iv) COSMIC-311 included the collection of EQ-5D-5L data up to the point of disease progression. The use of these data could have been explored in the CS, at least in sensitivity analyses. (v) The previous NICE appraisal of sorafenib and lenvatinib (TA535) applied treatment-specific utility values estimated using data from the DECISION trial. The Assessment Group's model applied a lower utility value for the TKIs versus BSC (progression-free utility value = 0.72 versus 0.80). Again, this could have been explored in sensitivity analyses in the CS.
What alternative approach has the EAG suggested?	The EAG believes that it may be reasonable to use the utility values reported by Fordham *et al.*, albeit based on the observed means rather than the higher values obtained from the adjusted regression model. This would provide consistency with several previous NICE appraisals of treatments for thyroid cancer. The EAG has undertaken additional sensitivity analyses which apply the utility values from the COSMIC-311 and DECISION trials.
What is the expected effect on the cost-effectiveness estimates?	Including QALY weighting, the EAG's error-corrected model suggests a deterministic ICER of £24,233 per QALY gained (EA1). Applying the observed mean estimates from Fordham *et al.* increases the ICER from £24,233 to £24,861 per QALY gained (EA3). Applying the utility value from COSMIC-311 increases the EAG's preferred ICER from £32,397 to £37,361 per QALY gained (ASA2a). Applying the utility value from the DECISION trial increases EAG's preferred ICER from £32,397 to £36,918 per QALY gained (ASA2b).
What additional evidence or analyses might help to resolve this key issue?	A judgement is required by the Appraisal Committee regarding whether it is more appropriate to apply utility values which are consistent with the target population (i.e., COSMIC-311) or those which are consistent with the majority of previous NICE appraisals of treatments for thyroid cancer (i.e., Fordham *et al.*). Further consideration should be given to whether the company's model adequately reflects the expected QALY losses associated with TKI-related toxicity whilst patients are progression-free and on treatment.

Issue 3: Issues relating to resource use and costs

Report section	5.3.5 (critical appraisal point [6])
Description of issue and why the EAG has identified it as important	The EAG has four concerns regarding the resource use and costing assumptions employed in the company's original model: • *Post-progression cabozantinib costs*. In COSMIC-311, patients in both treatment groups could receive open-label cabozantinib after progression. At CCO2, 6.5% of patients randomised to cabozantinib had received post-progression cabozantinib. In contrast, the company's economic model caps time to treatment discontinuation by PFS - this implies a stopping rule which was not employed in the trial. Given the experience of the COSMIC-311 trial and the company's intention for cabozantinib to be used in line with its licence, which permits continued treatment after progression for patients who are still clinically benefitting from treatment, the EAG believes that the costs of post-progression cabozantinib should be included in the economic analysis. • *Wastage costs*. The company's original model does not include any costs associated with drug wastage. In reality, patients who stop treatment due to progression or death before completing a full pack of treatment will incur some level of wastage. These costs should have been included. • *Monitoring cost assumptions*. The company's model assumes that patients receiving cabozantinib will undergo an electrocardiogram (ECG) once every 6 months. The EAG's clinical advisors suggested that patients would undergo ECGs more frequently. • *Concomitant medication costs.* The company's economic model does not include the costs of concomitant medications. The EAG believes that these should have been included.
What alternative approach has the EAG suggested?	The EAG's preferred analysis removes the company's modelled stopping rule and includes the costs of 7 days of drug wastage per patient. Additional exploratory analyses have also been undertaken to explore the impact of assuming more frequent ECGs for cabozantinib-treated patients and of excluding the costs of CT scans for BSC-treated patients. The EAG does not have sufficient information from COSMIC-311 to accurately estimate the costs of concomitant treatments.
What is the expected effect on the cost-effectiveness estimates?	When QALY weighting is included in the analysis, removing the stopping rule and modelling time to treatment discontinuation (TTD) using a Weibull distribution increases the ICER from the EAG's error-corrected model from £24,233 to £27,541 per QALY gained (EA4). Including drug wastage costs increases the EAG's error-corrected ICER from £24,233 to £24,686 per QALY gained (EA5). Applying alternative assumptions regarding resource use requirements for CT scans and ECGs has only a minor impact on the ICER (ASA4 and ASA5). Additional analyses undertaken by the EAG (not shown here) also indicate that the costs of concomitant therapies are unlikely to be a key driver of the ICER.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence or analysis is required to determine whether it is appropriate to include the costs of post-progression cabozantinib. Further follow-up would reduce uncertainty around TTD. Further analysis of data from COSMIC-311 could be used to estimate the costs of concomitant therapies.

1.6 Summary of EAG's preferred model and sensitivity analysis results

The results of the EAG's preferred model and additional sensitivity analyses are summarised in Table 2; results are presented with and without QALY weighting. Exploratory analysis 1 (EA1) reflects the EAG-corrected version of the company's model (deterministic). EA2-5 also include these corrections. EA6 is the EAG's preferred model. Additional sensitivity analyses (ASAs) use the EAG's preferred model (EA6) as a starting point.

Table 2: Summary of EAG's preferred model results

Scenario	DM	Inc. QALYs*	Inc. costs	ICER excluding QALY weighting	ICER including QALY weighting
Company's base case model					
Company's original base case (deterministic)	1.2	[redacted]	[redacted]	£28,148	**£23,456**
Company's original base case (probabilistic)	1.2	[redacted]	[redacted]	£27,169	**£22,641**
Company's revised base case (deterministic)	1.2	[redacted]	[redacted]	£24,347	**£20,289**
EAG preferred analysis†					
EA1: Correction of errors	1.2	[redacted]	[redacted]	£29,080	£24,233
EA2: Remove 5-year death assumption for BSC	1.2	[redacted]	[redacted]	£32,747	£27,289
EA3: Observed mean utility values from Fordham *et al.*	1.2	[redacted]	[redacted]	£29,834	£24,861
EA4: Stopping rule removed, TTD modelled using Weibull distribution	1.2	[redacted]	[redacted]	£33,050	£27,541
EA5: Inclusion of drug wastage costs	1.2	[redacted]	[redacted]	£29,623	£24,686
EA6a: EAG preferred analysis (deterministic)	1.2	[redacted]	[redacted]	£38,876	**£32,397**
EA6b: EAG preferred analysis (probabilistic)	1.2	[redacted]	[redacted]	£39,347	**£32,789**
EAG's additional sensitivity analyses					
ASA1a: Exponential OS with treatment effect waning at 3 years	1.2	[redacted]	[redacted]	£47,987	£39,989
ASA1b: Hybrid KM_{12} + exponential tail, constant HR	1.2	[redacted]	[redacted]	£40,675	£33,895
ASA1c: Hybrid KM_{12} + exponential tail, BSC hazard rate in both groups	1.2	[redacted]	[redacted]	£71,087	£59,240
ASA2a: COSMIC-311 utility value in progression-free state	1.2	[redacted]	[redacted]	£44,833	£37,361
ASA2b: DECISION trial utility values	1.2	[redacted]	[redacted]	£44,302	£36,918
ASA3: AE QALY losses doubled	1.2	[redacted]	[redacted]	£39,395	£32,829
ASA4: ECG costs doubled	1.2	[redacted]	[redacted]	£39,461	£32,884
ASA5: CT scan costs for BSC removed	1.2	[redacted]	[redacted]	£39,200	£32,667

Inc. - incremental.; DM - decision modifier; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; EA - exploratory analysis; ASA - additional sensitivity analysis; RDI - relative dose intensity; KM - Kaplan-Meier; BSC - best supportive care; AE - adverse event; EG - electrocardiogram; CT - computerised tomography
** Excluding QALY weighting*
† The EAG's analyses use the company's original model as a starting point

Modelling errors identified by the EAG are described in Section 5.3.5. For further details of the exploratory and sensitivity analyses undertaken by the EAG, see Section 5.5.

2. BACKGROUND

This chapter presents a brief summary and critique of the company's description of the disease (Section 2.1) and the company's overview of current treatment and their intended positioning of cabozantinib (Section 2.2). For completeness, additional information has been included by the External Assessment Group (EAG).

2.1 Company's description of the underlying health problem

Thyroid cancer (TC) is a rare type of cancer which accounts for around 1.2% of all malignancies in the UK. TC is caused by the growth of abnormal cells in the thyroid gland. There are four main types of TC: papillary, follicular, medullary and anaplastic. Data from Cancer Research UK indicate that between 2016 and 2018 there were 3,291 reported cases of TC and between 2017 and 2019 there were 341 deaths due to TC in England.[1] TC is more common in women than men, with women accounting for around 72% of all cases. Age-specific incidence rates peak at around age 45-49 years for women and at around age 70-75 years for men. Overall, the median age of diagnosis of TC is between 45 and 49 years.[1]

Differentiated thyroid cancer (DTC) is the most common form of TC, accounting for an estimated 90-95% of all cases.[2, 3] DTC cells have a similar appearance to normal thyroid cells and do not spread as quickly as undifferentiated cancer cells. DTC includes different subtypes of TC, including papillary thyroid carcinoma (PTC), follicular thyroid carcinoma (FTC) and Hürthle cell carcinoma. Of these, PTC is the most common subtype, which accounts for an estimated 83-86% of cases of DTC.[4, 5] Survival in DTC is strongly related to stage at diagnosis, with 1-year survival estimates ranging from 99% for Stages 1-3 (where the disease is localised to the thyroid) to 77% for Stage 4 (where the disease has spread beyond the thyroid).

Treatment of DTC typically involves surgery (usually thyroidectomy) which is usually used with curative intent. Subsequently, patients typically receive radioactive iodine (RAI) to destroy any cancerous cells not removed by surgery and those that have spread beyond the thyroid. However, it has been reported that between 5% and 15% of patients are refractory to RAI.[6] For patients who are RAI-refractory with locally advanced or metastatic DTC, first-line treatment usually involves the use of systemic therapy using a vascular endothelial growth factor receptor (VEGFR) tyrosine kinase inhibitor (TKI) – either sorafenib or lenvatinib. The population considered in this appraisal relates to patients with locally advanced or metastatic DTC, whose disease is refractory to, or who are unsuitable for RAI, and whose disease has progressed during or after prior systemic therapy. As noted in the company's submission (CS),[7] the expected survival for patients with RAI-refractory DTC is uncertain and is dependent on the availability of systemic therapies and the patient's prognosis. For RAI-refractory

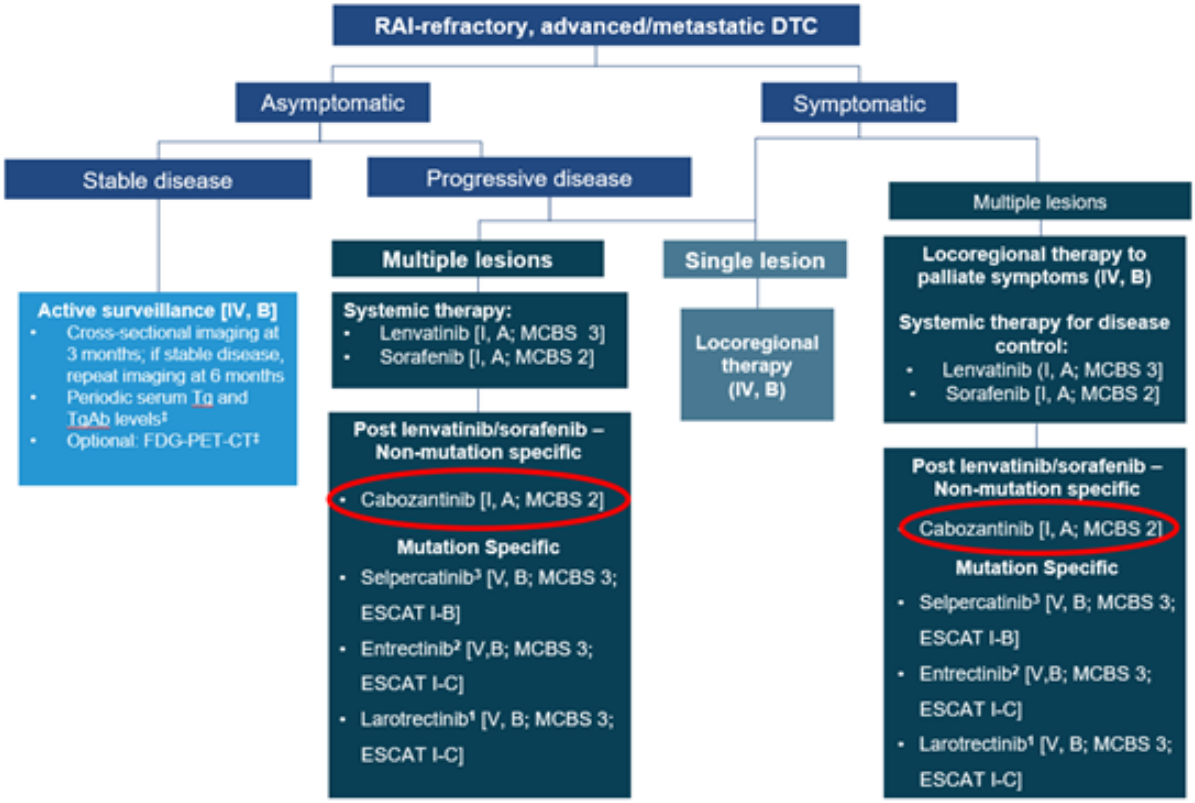
DTC, studies have reported 5-year, 10-year and 15-year survival rates of 66%, 10% and 6%, respectively.[8, 9] Expected survival is markedly worse in patients who have progressed on first-line systemic treatment: based on real-world data, the CS cites estimates of median survival for patients who have not received salvage therapy after progressing from a single agent TKI ranging from 10 months to 22 months.[10, 11] The CS does not discuss the impact of progressive RAI-refractory DTC on health-related quality of life (HRQoL). However, previous Technology Appraisal (TA) guidance documents published by the National Institute for Health and Care Excellence (NICE) have highlighted negative HRQoL impacts resulting from pain, fatigue, difficulty in carrying out daily activities and detrimental effects on emotional and mental health.[12, 13]

2.2 Company's overview of current service provision

2.2.1 Current treatment pathway for RAI-refractory DTC

The company's view of the current treatment pathway and proposed positioning of cabozantinib is reproduced in Figure 1. Current NICE recommendations for therapies used to treat patients with advanced or metastatic DTC are summarised in Table 3.

Figure 1: Company's view of the treatment pathway for RAI-refractory DTC and proposed positioning of cabozantinib, adapted from ESMO and NICE recommendations (reproduced from CS, Figure 2)



RAI - radioactive iodine; DTC - differentiated thyroid cancer; ESMO - European Society for Medical Oncology; FDG-PET-CT - [^{18}F]2-fluoro-2-deoxy-D-glucose-positron emission tomography-computed tomography; Tg - thyroglobulin; TgAb - serum thyroglobulin antibody; ESCAT - ESMO Scale for Clinical Actionability of Molecular Targets; MCBS - Magnitude of Clinical Benefit Scale

Table 3: Current NICE recommendations for treatments for advanced or metastatic DTC

NICE TA	NICE recommendation
TA535 - lenvatinib and sorafenib (2018)[12]	Lenvatinib and sorafenib are recommended as options for treating progressive, locally advanced or metastatic DTC (papillary, follicular or Hürthle cell) in adults whose disease does not respond to RAI, only if: • they have not had a TKI before or • they have had to stop taking a TKI within 3 months of starting it because of toxicity (specifically, toxicity that cannot be managed by dose delay or dose modification). Lenvatinib and sorafenib are recommended only if the companies provide them according to the commercial arrangements.
TA630 - larotrectinib (2020)[14]	Larotrectinib is recommended for use within the CDF as an option for treating NTRK fusion positive solid tumours in adults and children if: • the disease is locally advanced or metastatic or surgery could cause severe health problems and • they have no satisfactory treatment options. It is recommended only if the conditions in the MAA are followed.
TA644 - entrectinib (2020)[15]	Entrectinib is recommended for use within the CDF as an option for treating NTRK fusion-positive solid tumours in adults and children 12 years and older if: • the disease is locally advanced or metastatic or surgery could cause severe health problems and • they have not had an NTRK inhibitor before and • they have no satisfactory treatment options. It is recommended only if the conditions in the MAA for entrectinib are followed.
TA742 - selpercatinib (2021)[13]	Selpercatinib is recommended for use within the CDF, as an option for treating: • advanced RET fusion-positive TC in adults who need systemic therapy after sorafenib or lenvatinib • advanced RET-mutant MTC in people 12 years and older who need systemic therapy after cabozantinib or vandetanib. It is recommended only if the conditions in the MAA are followed.

TA - technology appraisal; NICE - National Institute for Health and Care Excellence; DTC - differentiated thyroid cancer; RAI - radioactive iodine; TKI - tyrosine kinase inhibitor; NTRK - neurotrophic tyrosine receptor kinase; CDF - Cancer Drugs Fund; RET - rearranged during transfection; TC - thyroid cancer; MTC - medullary thyroid cancer; MAA - Managed Access Agreement

Current NICE-recommended systemic first-line treatments for RAI-refractory advanced or metastatic DTC include two TKIs - lenvatinib and sorafenib (TA535). The recommendations for these treatments are restricted to patients who have not previously received treatment with a TKI before, although patients are able to switch from sorafenib to lenvatinib, or *vice versa*, within 3 months of starting treatment if toxicity occurs. There is currently no routinely commissioned NICE-recommended second-line treatment for patients who progress on first-line therapy. For patients who have progressed on systemic therapy and discontinued treatment, the only remaining option is best supportive care (BSC), which typically comprises thyroid stimulating hormone (TSH) suppression (e.g., using levothyroxine given indefinitely) and ongoing imaging, with palliative radiotherapy and symptom relief where necessary. Recent guidelines published by the European Society for Medical Oncology (ESMO)

mention both cabozantinib and lenvatinib as potential second-line treatments but state that there is uncertainty in that the optimal sequence cannot be determined based on currently available evidence.[16]

NICE has also issued positive recommendations for selpercatinib for the treatment of advanced rearranged during transfection (RET) fusion-positive TC (TA742)[13] and for larotrectinib and entrectinib for the treatment of neurotrophic tyrosine receptor kinase (NTRK) fusion-positive solid tumours which may include DTC (TA630 and TA644).[14, 15] Selpercatinib is recommended for patients with advanced RET fusion-positive TC after lenvatinib or sorafenib. Entrectinib and larotrectinib are recommended for patients with NTRK fusion positive solid tumours (including thyroid cancer), with the manufacturers of both products positioning these treatments as last-line therapies. Selpercatinib, entrectinib and larotrectinib are all currently recommended only for use in the Cancer Drugs Fund (CDF) and do not form part of routine NHS commissioning.

2.2.2 Company's proposed positioning of cabozantinib

The company's proposed positioning for cabozantinib is in line with the licensed indication for cabozantinib, that is, in adult patients with locally advanced or metastatic DTC, refractory to or not eligible to receive RAI who have progressed during or after prior systemic therapy. The company's clarification response[17] (question A14) states that this could be as second- or later-line treatment following prior systemic therapy.

2.2.3 EAG clinical advisors' views

The EAG's clinical advisors agreed with the company's description of the disease and their proposed positioning of cabozantinib. The clinical advisors commented that whilst more women are diagnosed with DTC than men, in the metastatic setting, the proportions of men and women are similar. The advisors commented that there are some patients in whom they would not consider offering cabozantinib. One of the clinical advisors stated that they would not offer cabozantinib to patients who have a prolonged QT interval, but highlighted that this reflects a minority of patients. The EAG notes that the Summary of Product Characteristics (SmPC) for lenvatinib, sorafenib and cabozantinib each state warnings about the use of these treatments in patients with prolonged QT intervals. The second advisor commented that they would be concerned about offering cabozantinib to patients with poor performance status (PS≥2) and/or to frail elderly patients. One advisor further stated that elderly patients who are still fit would still be considered for treatment. The advisors agreed that current NICE recommendations only allow for the use of either lenvatinib or sorafenib as first-line therapy, except where patients switch TKI due to toxicity (within 3 months of starting treatment), and that cabozantinib would be used as second-line therapy in most patients.

3. CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

This chapter presents a summary and critique of the decision problem addressed by the CS.[7] A summary of the decision problem as outlined in the final NICE scope[18] and addressed in the CS is presented in Table 4. The EAG's critique of the decision problem addressed within the CS is presented in the subsequent sections.

Table 4: The decision problem (reproduced from CS, Table 1, with minor amendments and comments from the EAG)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from final NICE scope	EAG comments
Population	Adults with locally advanced or metastatic DTC, whose disease is refractory to, or who are unsuitable for RAI, and whose disease has progressed during or after prior systemic therapy.	Adults with locally advanced or metastatic DTC, whose disease is refractory to, or who are unsuitable for RAI, and whose disease has progressed during or after prior systemic therapy.	N/a	In line with the final NICE scope. Twenty-four percent of patients in COSMIC-311 had received two prior VEGFR-TKIs (sorafenib and lenvatinib). In usual practice in England, patients would only receive either sorafenib or lenvatinib (not both), except where patients switch TKI due to toxicity.
Intervention	Cabozantinib (Cabometyx®)	Cabozantinib (Cabometyx®)	N/a	In line with the final NICE scope. The company's proposed positioning includes treatment beyond progression, but these costs are excluded from the model. Cabozantinib is given alongside BSC.
Comparator(s)	BSC	BSC	As per the final scope, BSC is the comparator. There are no other treatments recommended post first-line systemic treatment for RAI refractory DTC patients by NICE, NHSE or ESMO. ESMO does state that 'cabozantinib and lenvatinib [are] two potential choices for second-line treatment of patients who progress on sorafenib'. However, as described earlier, the sequence of treatment should be determined on each patient's response and ESMO cannot create an optimal sequence for advanced/metastatic DTC due to limited current evidence.[16]	Clinical advisors consulted by the company indicated that lenvatinib is currently the first-line treatment of choice and some clinicians continue to offer this treatment after progression on this therapy. However, continued lenvatinib given post-progression is not included as a comparator in the final NICE scope or in the CS.

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from final NICE scope	EAG comments
Outcomes	Draft scope: • Overall survival • Progression-free survival • Response rate • Adverse effects of treatment • Health-related quality of life	Co-primary endpoints: • Objective response rate (confirmed per RECIST v1.1) • Progression-free survival • Additional endpoints • Overall survival • Adverse effects of treatment • Health-related quality of life	N/a	In line with the final NICE scope.
Impact of the technology beyond direct health benefits, and on the delivery of the specialised service*	No special considerations stated in the final scope.	No special considerations stated in the final scope.	N/a	Not relevant to NICE STAs
Special considerations including issues related to equity or equality	No special considerations stated in the final scope.	No special considerations stated in the final scope.	Further to the company's decision problem and final scope, we believe that special considerations should be made regarding the female prevalence of DTC. Females are much more likely to be diagnosed with thyroid cancer making up 72% of thyroid cancer cases in the UK. In England, the AS incidence rate for thyroid cancer in females is 8.7 and for male it is 3.6 per 100,000, respectively, a clear difference in the incidence between females and males.[1]	The population included in COSMIC-311, which is reflected in the company's model, includes a comparatively lower proportion of women (53%). The EAG's clinical advisors commented that in the metastatic setting, the proportions of men and women are similar.

NICE - National Institute for Health and Care Excellence; CS - company's submission; EAG - External Assessment Group; VEGFR-TKIs - vascular endothelial growth factor receptor tyrosine kinase inhibitor; BSC - best supportive care; RAI - radioactive iodine; DTC - differentiated thyroid cancer; NHSE - NHS England; ESMO - European Society for Medical Oncology; RECIST - Response Evaluation Criteria in Solid Tumours; STA - Single Technology Appraisal; AS - age-standardised; N/a - not applicable

** The EAG is unsure why this item has been included in the company's summary of the decision problem*

3.1 Population

The final NICE scope[18] specifies the relevant population as adults with locally advanced or metastatic DTC, whose disease is refractory to, or who are unsuitable for RAI, and whose disease has progressed during or after prior systemic therapy. The main clinical evidence for cabozantinib included in the CS[7] comes from the COSMIC-311 randomised controlled trial (RCT).[19] Patients enrolled in COSMIC-311 were individuals aged 16 years and older with RAI-refractory DTC (papillary or follicular and their variants) who had received previous lenvatinib or sorafenib and progressed during or after treatment with up to two VEGFR TKIs. Patients in the trial had an Eastern Cooperative Oncology Group (ECOG) PS of 0 or 1. The European Medicines Agency / Medicines and Healthcare products Regulatory Agency (EMA/MHRA) marketing authorisation for cabozantinib relates to *"adult patients with locally advanced or metastatic differentiated thyroid carcinoma (DTC), refractory or not eligible to RAI who have progressed during or after prior systemic therapy."*[20] The company's clarification response (question A14) states that the company is seeking a positive recommendation for cabozantinib in line with its full licensed indication, that is, as second- or later-line therapy following prior systemic therapy.

The company's clinical evidence and economic model reflect data from the full intention-to-treat (ITT) population of COSMIC-311, in which 76% of patients had received one prior TKI (sorafenib or lenvatinib) and the remainder had received both sorafenib and lenvatinib.[19] The EAG's clinical advisors stated that in England patients would only receive one prior TKI therapy (most likely lenvatinib). One of the advisors commented that patients who have had two prior TKIs may be a worse prognostic group and therefore outcomes from COSMIC-311 might have been improved slightly if only patients with one prior treatment were recruited. However, they also commented that if patients had discontinued previous treatments because of toxicity, there may be no difference between those who have received one prior treatment or two prior treatments. The second clinical advisor commented that patients with two prior treatments may have a greater burden of cumulative toxicity from previous treatments (particularly fatigue, skin toxicity and hypertension) which might affect how long patients can remain on cabozantinib; they were unsure about the extent to which the number of prior therapies might impact on clinical outcomes.

3.2 Intervention

The intervention described in the CS[7] is consistent with the final NICE scope.[18] The intervention under consideration is cabozantinib (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.[20]

A full marketing authorisation for cabozantinib in the DTC indication was issued by the EMA in April 2022. The MHRA granted a Type II extension for the use of cabozantinib for the treatment of DTC in

May 2022.[7] Cabozantinib is administered orally in tablet form at a recommended dose of 60mg once daily. The SmPC[20] states that treatment interruptions and/or dose reductions may be required in the event of adverse drug reactions; the dose should be reduced initially to 40mg daily and then to 20mg daily. The list price per pack of 30 x 60mg cabozantinib tablets (30 days' supply) is £5,143.00. A confidential Patient Access Scheme (PAS) discount is available for cabozantinib which takes the form of a simple price discount of [redacted]. The price per pack of cabozantinib including the PAS is [redacted]. The list price and PAS discount per pack of cabozantinib at the 40mg and 20mg doses is the same as that for the 60mg dose.

The EAG notes that within the COSMIC-311 trial,[19] patients randomised to the cabozantinib group were allowed to continue to receive open-label cabozantinib after progression. At clinical cut-off (CCO) 2 (8th February 2021), 6.5% of patients randomised to cabozantinib had received post-progression cabozantinib. The company's clarification response[17] (question A15) states that the company is seeking a positive approval for cabozantinib use which is in line with the SmPC,[20] i.e., treatment should be continued until the patient is no longer clinically benefitting from therapy or until unacceptable toxicity occurs. In contrast, the company's economic model implicitly assumes that patients will stop treatment at the point of disease progression, but does not include any adjustment of OS for post-progression treatment in the cabozantinib group and the costs of post-progression cabozantinib are excluded. The EAG's clinical advisors commented that they would want to use cabozantinib in the same way that they use current first-line TKIs, with treatment being discontinued at progression, but noted that other clinicians may wish to continue treatment for longer if the patient is still deriving clinical benefit. This issue is discussed further in Section 5.3.5.

3.3 Comparators

The final NICE scope[18] lists a single comparator – best supportive care (BSC). The COSMIC-311 trial[19] was placebo-controlled and the comparator considered within the CS[7] and the company's economic model is BSC alone. [redacted]Continued lenvatinib given post-progression is not included in the company's economic model and was not listed as a comparator in the final NICE scope. The EAG considers that it is unlikely that sufficient evidence exists to inform a reliable comparison between cabozantinib and continued lenvatinib given post-progression.

3.4 Outcomes

The following outcomes are listed in the final NICE scope:[18]

- Overall survival (OS)
- Progression-free survival (PFS)
- Response rate
- Adverse effects of treatment
- Health-related quality of life (HRQoL).

The CS[7] reports on all of these outcomes for the ITT population of COSMIC-311.[19] The company's economic model is informed by data on PFS, OS, and adverse events (AEs). HRQoL data collected in COSMIC-311[19] are not used in the model; instead the utility values used in the model have been derived from an external valuation study[22, 23] (see Section 5.2.4).

3.5 Other relevant factors

The CS[7] reports that women are much more likely to be diagnosed with TC and cites current estimates from Cancer Research UK which suggest that women represent 72% of all cases of TC.[1] The CS states that a positive recommendation for cabozantinib in the treatment of RAI-refractory DTC will reduce health inequalities for female TC patients.

The EAG notes that the estimates quoted from Cancer Research UK represent all thyroid cases and do not specifically reflect the target population for cabozantinib. The EAG also notes that whilst TC incidence is higher in women, studies have indicated that men have a worse prognosis than women.[24] The EAG's clinical advisors commented that whilst TC disproportionately affects women, the proportions of men and women seen in the advanced/metastatic DTC setting are similar.

4. CLINICAL EFFECTIVENESS

The clinical evidence contained in the CS[7] is comprised of:

- A systematic literature review (SLR) of clinical evidence for cabozantinib for treating adult patients with RAI-refractory DTC receiving second-line or third-line treatment
- Summary and results for the COSMIC-311 trial of cabozantinib.

This chapter summarises and critiques the company's review methods and clinical effectiveness data. Full details are presented in the Section B.2 of the CS[7] and CS Appendices D, E and F.[25]

4.1 Critique of the methods of review(s)

The clinical evidence presented in the CS[7] was informed by an SLR of studies assessing the clinical efficacy and safety of cabozantinib in adult patients with RAI-refractory DTC receiving second- or third-line treatment, who have been previously treated with sorafenib and/or lenvatinib (CS Appendix D.1.1,[25] Table 2). The primary clinical evidence detailed in the CS comes from COSMIC-311 (XL184-311; NCT03690388) – an international Phase III, multi-centre randomised controlled trial (RCT). Eight publications relating to this trial were identified by the SLR and update search (CS, Appendix D.1.2, Table 7).[19, 26-32] COSMIC-311 compared cabozantinib plus BSC with placebo plus BSC. In the trial, BSC included: analgesia; antibiotics for infections; transfusions for anaemia; nutritional support, and psychological support with medication or counselling as appropriate.[19]

The CS[7] identified BSC as the principal comparator for cabozantinib for the DTC indication in England because there is currently no licensed treatment for this population who have previously progressed on sorafenib and/or lenvatinib (CS, Sections A.8 and B.2.2). Given the availability of a head-to-head Phase III RCT comparing cabozantinib with the BSC, and the absence of any trials of other licensed treatments for this indication, the CS argues that an indirect treatment comparison (ITC) and network meta-analysis (NMA) was not necessary (CS, Sections B.2.2 and B.2.9).

The safety evidence reported in the CS[7] comprised a narrative summary of data from the COSMIC-311 trial (CS, Sections B.2.10).

4.1.1 Searches

Appendix D of the CS[25] reports an SLR of clinical effectiveness evidence in RAI-refractory second- or third-line DTC. The company's searches covered all of the core databases recommended by NICE: MEDLINE (including In-Process and Epub-ahead-of-print citations); EMBASE; the Cochrane Library; and relevant conferences and trial registers.

The search strategies are reported in full in Tables 2-4 of CS Appendix D[25] (pages 6-13). However, as the company searched MEDLINE and EMBASE together via the ProQuest platform, it was not possible for the EAG to replicate them as executed. The EAG generally advises against multi-file searching when conducting a systematic review – whilst there are clearly efficiency savings, the way in which a string designed for one database is interpreted by another can vary between different search interfaces with unforeseen effects. The company maintains that their approach will not have missed any relevant publications, but as the EAG does not have access to the ProQuest platform, it can neither disprove this assertion nor verify it with any certainty. However, the searches are well designed, combining appropriate subject indexing and free text terms for the population of interest with study filters based on those developed by the Scottish Intercollegiate Guidelines Network (SIGN). Whilst these filters have not been formally validated, they are widely used in the evidence synthesis community and they are generally agreed to be unlikely to miss study types eligible for inclusion.

The EAG notes that the searches were conducted 27th of September 2021. During the clarification process, the EAG asked the company to explain why a full update search had not been completed for the period October 2021-September 2022 (see clarification response,[17] question A5). The company's response states that this was due to scheduling issues and that such a search was considered "impractical." Instead, the company conducted a "targeted search using internal Ipsen databases", justifying this on the basis that *"Ipsen are following this disease space closely"* and that given the company's focus on this indication, their own database would include all likely publications from that period. Whilst this explanation seems plausible, the EAG notes that this does not constitute a systematic update search, and the company's approach overlooks the fact that systematic searches for the purpose of SLRs are not only about finding the relevant evidence but also about demonstrating that this has been done assiduously so that the reader knows that nothing has been omitted either deliberately or inadvertently. Regrettably, the lack of transparency about the update process means that the EAG can be less confident about the comprehensive retrieval of evidence published since the completion of the main searches; however, it is reasonable to believe that their knowledge of studies of their own product (cabozantinib) is complete.

4.1.2 Inclusion criteria for the SLR

The inclusion and exclusion criteria for the SLR are reported in Table 5. These criteria differ from the final NICE scope[18] with respect to the population, the intervention and the comparators (see Table 4). Unlike the NICE scope, which refers to adult patients only, the eligibility criteria for the SLR included both adult patients and adolescent patients aged ≥12 years. The intervention detailed in the NICE scope was limited to cabozantinib and the comparator was limited to BSC alone. However, the company's SLR included more than 20 potential interventions (including cabozantinib) and comparators (including BSC). The clinical effectiveness section of the CS[7] only reports studies relevant to the decision problem

and the NICE scope, principally the COSMIC-311 trial.[19] The CS also identified two potentially relevant Phase II trials of cabozantinib in the required population or a similar population (CS, Section B.2.8). These trials were only cited in the section on meta-analysis in the CS (Section B.2.8), and both studies were excluded from the SLR. The CS states that one trial was excluded because the population was correct but related to first-line treatment only (no previous treatment with sorafenib or lenvatinib) (NCT02041260, reported in Brose 2014[33]). A second single-arm study was excluded because there was potential for patients to dose-escalate cabozantinib treatment to 80mg (rather than remain on the licensed dose of 60mg) (NCT01811212, reported in Cabanillas *et al.*, 2017[34]). The EAG notes that only 4 out of 25 (16%) patients received cabozantinib at the higher dose in this study and the EAG questioned its exclusion based on the higher dose, including noting that dose was not a pre-specified exclusion criterion for the clinical effectiveness review (see clarification response,[17] questions A7 and A9). The company's clarification response states that the trial was not considered to be relevant to the NICE decision problem, and it permitted a dose which was different from the licensed regimen. However, in response to the EAG's clarification request, the company considered this trial to be potentially relevant and briefly summarised the trial and its findings in their clarification response (question A9).

Given the case presented in Sections B.2.2 and B.2.9 of the CS[7] for not conducting any ITCs, which were applied *a priori*, the EAG questioned the rationale for the initial inclusion of so many interventions and comparators in the SLR, the vast majority of which were considered not relevant to the decision problem and were not included in clinical effectiveness section of the CS. The company's clarification response[17] (question A7) acknowledges that the SLR inclusion criteria are broader than the NICE decision problem, but argues that the SLR was conducted *"to try and satisfy the needs of multiple HTA countries"*, to which the criteria of the NICE decision problem were then applied, and that this was not made clear in the CS.

Table 5: Inclusion and exclusion criteria for the SLR (adapted from CS Appendix D.1.1, Table 1)

	Inclusion criteria	Exclusion criteria
Population (P)	• Patients with RAI-refractory DTC • Locally advanced/metastatic disease • Adolescent and adult patients (≥12 years) • Received ≥ prior MKI /TKI or VEGFR-targeted therapy	• DTC patients with early-stage or local regional disease • Paediatrics (<12 years of age)
Intervention (I)	• Cabozantinib • Dabrafenib • Sorafenib • Trametinib • Lenvatinib • Larotrectinib • Sunitinib • Nivolumab • Selpercatinib • Ipilimumab • Pralsetinib • Sirolimus • Pazopanib • Everolimus • Vandetanib • Regorafenib • Nintedanib • Avelumab • Vemurafenib • Temsirolimus • Apatinib • Atezolizumab • Pembrolizumab • Cobimetinib • Durvalumab • Any other systematic therapy evaluated in clinical practice	Any treatment that is not listed in the inclusion criteria
Comparators (C)	• Any intervention matching the intervention criteria • Placebo • Best supportive care	Any treatment that is not listed in the inclusion criteria
Outcomes (O)*	**Efficacy** • OS • PFS • Tumor response (BOCR, ORR, CR, PR, SD, PD) • DoR • Disease stabilisation rate • Time to response • TTD **Safety** • Incidence grade ≥3, SAEs, TRAEs and TEAEs • Discontinuation rates due to AEs **HRQoL/PRO** • EQ-5D (3L and 5L) • HUI • SF-36 (incl. variations) • TTO • SG • EORTC-QLQ-C30 • EORTC-QLQ-THY34 • THYCA-QoL • FACIT/FACT instruments • ThyPRO	Studies that do not report any of the outcomes of interest specified in the inclusion criteria
Study Design (S)	• RCTs • Non-randomized prospective interventional trials • Observational studies (prospective or retrospective) • Systematic literature reviews and meta-analyses	• Preclinical studies • Editorials • Commentaries • Erratum • Letters
Time restriction	• Peer-reviewed publications: no time restriction • Conference proceedings: 2015 or later	• Peer reviewed publications: no time restriction • Conference proceedings: <2015
Language	• No language restrictions	• N/a

RAI - radioactive iodine; DTC - differentiated thyroid cancer; MKI - multikinase inhibitor; TKI - tyrosine kinase inhibitor; VEGFR - vascular endothelial growth factor receptor; RW - real-world; BOCR - best overall confirmed response; ORR - objective response rate; CR - complete response; PR - partial response; SD - stable disease; PD - progressed disease; DoR - duration of response; TTD - time to treatment discontinuation; SAE - serious adverse event; TRAE - treatment-related adverse event; TEAE - treatment-emergent adverse event; AE - adverse event; EQ-5D - Euroqol 5-Dimensions; HUI - Health Utilities Index; SF-36 - Short Form 36; TTO - time trade-off; SG - standard gamble; EORTC-QLQ-C30/ThY34 - European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core/Thyroid; FACIT - Functional Assessment of Chronic Illness Therapy; FACT - Functional Assessment of Cancer Therapy; ThyPRO - Thyroid-Related Quality of Life; RCT - randomised controlled trial; N/a – not applicable

**Reported for least 90% of patients or as separate subgroup.*

The SLR criteria included the key effectiveness outcomes from the final NICE scope:[18] OS, PFS, response rates, HRQoL and safety outcomes (CS, Section B.1.1 and CS Appendix D.1.1 Table 1).[7, 25]

4.1.3 Critique of study selection, data extraction and quality assessment

CS Appendix D.1.1[25] reports that, for all citations, both the title/abstract and full-text screening stages of study selection were undertaken independently by two reviewers, and any discrepancies were reconciled by a third independent reviewer. The EAG considers independent study selection by two or more reviewers, as conducted here, to be best practice in systematic reviewing. The EAG notes that publications relating to the principal trial, COSMIC-311,[19] identified in an update search were not included in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart detailing the results of the study selection process; this was corrected in the company's clarification response[17] (question A6).

The company's data extraction methods are reported in CS Appendix D.1.1.[25] Data extracted from the included studies are presented in Sections B.2.3-2.7 and 2.10 of the CS and in CS Appendix D1.2. The process was undertaken independently by two reviewers, and any discrepancies were reconciled by a third independent reviewer. The EAG considers independent data extraction by two or more reviewers, as conducted here, to be best practice in systematic reviewing.

No details were provided regarding the process followed in the conduct of quality assessment of studies included the clinical effectiveness review. This information was later provided in the company response to clarification question A10.[17] The process was undertaken by one reviewer, checked by a second, and any discrepancies were reconciled by a third reviewer. The EAG considers independent risk of bias/quality assessment by two or more reviewers to be best practice in systematic reviewing.

4.1.4 Results of the company's SLR

The clinical SLR presented in the CS[7] identified one Phase III trial of cabozantinib that was relevant to the decision problem: COSMIC-311 (XL184-311; NCT03690388).[19] This study forms the key evidence for clinical effectiveness and safety of cabozantinib within the CS. Eight publications were identified and listed for this study.[19, 26-32] The EAG believes that no additional relevant published Phase III trials of cabozantinib in adult patients with RAI-refractory DTC receiving second- or third-line treatment that could have provided data on safety and efficacy have been omitted from the CS. However, a second trial that did satisfy the criteria was identified in the CS[7] but was excluded based on a proportion of patients who experienced dose escalation of cabozantinib above the licensed dose (80mg rather than 60mg).[34] This exclusion was queried by the EAG and the company subsequently agreed to the inclusion of this trial and its data (clarification response,[17] question A9). This trial is briefly described in Section 4.5.

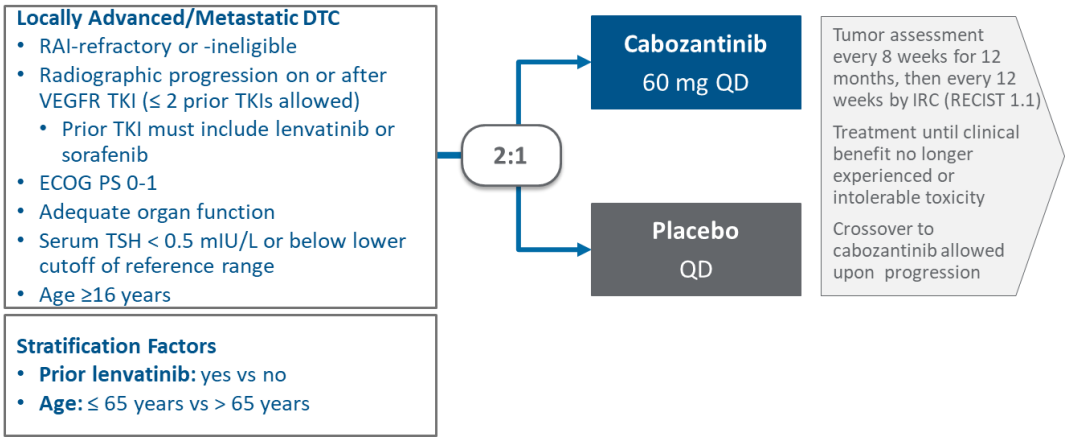
4.2 Characteristics of the COSMIC-311 study of cabozantinib

4.2.1 Study design: COSMIC-311

COSMIC-311 is a Phase III, randomised, international, multi-centre, blinded, parallel-arm trial initiated in May 2018 and conducted in 164 centres across 25 countries (NCT03690388). The primary completion date was August 2020, but the final completion date is listed as December 2022 (NCT03690388). Overall, 227 patients were screened and 187 adult patients with RAI-refractory DTC receiving second- or third-line treatment who satisfied all eligibility criteria were randomised.[19]

Details of study location, treatments, inclusion and exclusion criteria, prohibited concomitant medications and relevant outcomes are reported in Table 6. Patients were initially selected based on the eligibility criteria described in Table 6 and were assigned in a 2:1 ratio to either the intervention arm (cabozantinib plus BSC) or the control arm (placebo plus BSC). Randomisation was stratified by previous lenvatinib (yes vs. no) and age (≤65 years vs. >65 years).[19] The patient cohorts assessed in the clinical effectiveness review are presented in Figure 2.

Figure 2: Overview of trial design for COSMIC-311 (reproduced from CS, Section B.2.3.1, Figure 3)



ECOG - Eastern Cooperative Oncology Group; IRC - Independent Radiology Committee; QD - once a day; RAI - radioactive iodine; RECIST - Response Evaluation Criteria in Solid Tumours; TKI - tyrosine kinase inhibitor; TSH - thyroid-stimulating hormone; VEGFR - vascular endothelial growth factor receptor
Source: XL184-311 CSR (30th April 2020)

Patients received self-administered 60mg cabozantinib daily or a placebo equivalent. Patients continued treatment in either arm until disease progression as confirmed by Response Evaluation Criteria in Solid Tumours (RECIST) version 1.1 or unacceptable toxicity; other reasons for discontinuation included patient decision, non-compliance or pregnancy.[19] Patients who were unmasked at radiographic progression and found to be in the placebo group could cross over, if eligible, to receive open-label cabozantinib 60mg daily. Patients who were unmasked at radiographic progression and found to be in the cabozantinib group could also transition to open-label cabozantinib (60mg daily) as long as the investigator deemed them to be deriving clinical benefit.

Table 6: Summary of trial methodology of COSMIC-311 (reproduced from CS, Table 4)

Trial	COSMIC-311
Trial design	Phase 3, randomised, multicentre, double-blind, 2:1 controlled study of cabozantinib versus placebo in patients with RAI-refractory DTC who have received prior lenvatinib or sorafenib treatment.

Eligibility criteria for participants	**Key Inclusion criteria:**
	• Histologically or cytologically confirmed diagnosis of DTC, including the following subtypes:
	  o PTC including histological variants of PTC
	  o FTC including histological variants of FTC
	• Measurable disease according to RECIST 1.1 on CT/MRI performed within 28 days prior to randomisation
	• Must have been previously treated with or deemed ineligible for treatment with Iodine-131 for DTC
	• Patients must have received at least one prior VEGFR-targeting TKI therapy of either lenvatinib or sorafenib and must have had radiographic progression during treatment or within 6 months after the most recent dose of the VEGFR inhibitor (up to two prior therapies were allowed including, but not limited to, lenvatinib and sorafenib)
	• Must have experienced documented radiographic progression per RECIST 1.1 per the Investigator during or following treatment with a VEGFR-targeting TKI prior to starting the next anticancer therapy (which may have been treatment in COSMIC-311)
	• Age – 16 years and older (Adult, Older Adult)
	• ECOG PS of 0 or 1

Exclusion Criteria for participants	**Key Exclusion criteria:**
	• Prior treatment with any of the following:
	  o Cabozantinib
	  o Selective small-molecule BRAF kinase inhibitor (e.g., vemurafenib, dabrafenib)
	  o More than 2 VEGFR-targeting TKI agents (e.g., lenvatinib, sorafenib, sunitinib, pazopanib, axitinib, vandetanib)
	  o More than 1 immune checkpoint inhibitor therapy (e.g., PD-1 or PD-L1 targeting agent)
	  o More than 1 systemic chemotherapy regimen (given as single agent or in combination with another chemotherapy agent)
	• Receipt of any type of small molecule kinase inhibitor within 2 weeks or 5 half-lives of the agent, whichever was longer, before randomisation
	• Receipt of any type of anticancer antibody or systemic chemotherapy within 4 weeks before randomisation
	• Receipt of radiation therapy for bone metastasis within 2 weeks or any other radiation therapy within 4 weeks before randomisation
	• Subjects with clinically relevant ongoing from prior radiation therapy that had not completely resolved were not eligible
	All inclusion and exclusion criteria are listed in Appendix C.

Settings and locations where the data were collected	A total of 258 subjects were randomised in 161 unique sites by 174 principal investigators in 25 countries in Asia, North America, Europe, and the rest of the world. These included: • Europe: Austria, Belgium, Croatia, Czech Republic, France, Germany, Hungary, Italy, Netherlands, Poland, Romania, Spain, United Kingdom • North America: United States of America and Canada • Asia: Hong Kong, Republic of Korea, Taiwan, Thailand • Rest of the world: Argentina, Australia, Brazil, Israel, Mexico, Russia,

Trial	COSMIC-311
Trial	**COSMIC-311**
Trial drugs	• Experimental Arm: Cabozantinib 60 mg tablet once daily ○ Two dose reductions in decrements of 20 mg was permitted to manage or prevent AE or toxicity • Comparator Arm: Matched placebo
Permitted and disallowed concomitant medication	**Allowed concomitant medication** • Prophylactic antiemetics and antidiarrheal medications in line with standard clinical practice • Granulocyte colony-stimulating factors per ASCO or ESMO guidelines • Bisphosphonates or denosumab for the control of bone loss or hypercalcemia if the benefit per the Investigator's discretion • Transfusions and hormone replacement (including TSH-suppressive thyroid hormone therapy) • Prophylactic individualised anticoagulation therapy with low dose low molecular weight (LMWH) heparins for supportive treatment per the Investigator's discretion. LMWH use at first dose should only be used if the subject had no evidence of brain metastasis, had been on stable dose of LMWH for a least six weeks prior, and had no complications from a thromboembolic event or the anticoagulation regimen. Therapeutic doses of oral anticoagulants (e.g., warfarin or other coumarin-related agents) were not allowed after randomisation until study treatment was permanently discontinued **Prohibited Therapies** • Any investigational agent or investigational medical device • Any systemic NPACT (e.g., chemotherapy, immunotherapy, radionuclides, drugs, or herbal products used specifically for the treatment of DTC). • Therapeutic doses of oral anticoagulants. • Local anticancer treatment including palliative radiation, ablation, embolisation or surgery impacting on tumour lesions were only to be performed until radiographic progression was confirmed per RECIST 1.1. • Erythropoietic-stimulating agents prohibited due to the increased risk of tumour recurrence. • Concomitant medications that prolong the QTc interval were to be avoided until subjects discontinue treatment. • Chronic coadministration of strong CYP 3A4 inducers due to potential to decrease exposure to cabozantinib. • Coadministration of strong CYP3A4 inhibitors and other drugs that inhibit CYP3A4 was to be avoided because these drugs had the potential to increase exposure (AUC) to cabozantinib

AE - adverse event; AUC - area under the curve; ASCO - American Society of Clinical Oncology; CT - computed tomography; CYP - cytochrome P450; DTC - differentiated thyroid cancer; ECOG - Eastern Cooperative Oncology Group; ESMO - European Society of Medical Oncology; FTC - follicular thyroid carcinoma; LMWH - low molecular weight heparin; MRI - magnetic resonance imaging; NPACT - non-protocol anticancer therapy; PD-1 - programmed cell death-1; PD-L1 - programmed cell death ligand 1; PS - performance status; PTC - papillary thyroid carcinoma; QTc - corrected QT interval; RAI - radioactive iodine; RECIST - Response Evaluation Criteria in Solid Tumours; TKI - tyrosine kinase inhibitor; VEGFR - vascular endothelial growth factor receptor

Source: XL184-311 CSR (30th April 2020) and Brose et al., 2021[19]

4.2.2 Quality assessment of COSMIC-311

The company's quality assessment of COSMIC-311 was undertaken using the Centre for Reviews and Dissemination (CRD) checklist for RCTs[35] (as per recommendations in the NICE user guide). The findings of this quality assessment are reported in Section B.2.5 of the CS;[7] these are reproduced in Table 7 together with the EAG's judgements. However, the assessments depend on whether they relate to the original masked phase of the trial, or the unmasked/crossover phase of the trial. For this reason, the EAG has included assessments for both phases. The EAG agrees with the company's responses to some of the checklist criteria: randomisation was conducted appropriately; treatment allocation concealment was adequate pre-crossover; there were no unexpected imbalances between drop-outs; there is no evidence of selective outcome reporting, and an appropriate intention-to-treat (ITT) analysis was used.

However, in the crossover phase of the trial, after unmasking for patients who progressed, and their assignment to possible open-label cabozantinib treatment, allocation concealment was judged by the EAG not to be adequate - this affected the safety and response outcomes. The EAG agrees that the two arms were balanced in terms of most known prognostic factors before the crossover phase, with the exception of tumour volume/burden, which was not controlled for.[36, 37] However, the balance between participants in each arm was compromised by the crossover phase of the trial. All assessments were blinded in the pre-crossover phase of the trial, but the response and safety assessments were not blinded in the crossover phase (CS,[7] Section B.2.3.2).

The EAG also conducted a quality assessment using the Cochrane Risk of Bias tool (version 2),[38] which is the international standard for the quality assessment of RCTs. This assessment is presented in Table 8. The risk of bias due to missing data and selective reporting was judged to be low. The risks of performance bias and outcome assessment bias were judged to be low or associated with some concerns due to the failure to control for the prognostic factor of tumour volume/burden and the unblinded assessment of some outcomes after crossover.

Table 7: Quality assessment of COSMIC-311 including the EAG's critique (based on data presented in Brose *et al.*, 2021 and CS, Section B)

Study question	Response (yes/no/not clear/NA) CS / EAG	How is the question addressed in the study? EAG
Was randomisation carried out appropriately?	Yes / Yes	*Randomisation was stratified by previous Lenvatinib treatment (yes vs no) and age (≤65 vs >65 years). The randomisation scheme used stratified permuted blocks of block size six and study treatment was centrally assigned through an interactive voice–web response system. Generation of the randomisation schedule was assigned to a clinical research organisation who maintained an unmasked team independent from the study. The live schedule, generated by the clinical research organisation, was uploaded to a secured server for the interactive response technology vendor who was responsible for interactive voice–web response services. Study personnel did not have access to the live schedule, the master list of blocks or block sizes, until authorised and documented unmasking (April 16, 2021)*[19]
Was the concealment of treatment allocation adequate?	Yes / Yes and No	Open-label treatment with cabozantinib was permitted for the following eligible patients: patients in the cabozantinib arm who progressed; patients in the placebo arm who progressed. OS, ORR and safety data were reported for these patients.[19]
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes / Yes and not clear	Yes, the arms were balanced at baseline for known prognostic factors such as gender, but no information was provided on the prognostic factor of tumour volume/burden[36, 37]
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes /	*Unique drug pack numbers were preprinted onto each bottle or package and assigned to the patient by the interactive voice–web response system to ensure patients, investigators, site staff, and the study sponsor remained masked to treatment assignment. Investigators could request that patients be unmasked at the time of radiographic progression confirmed by blinded independent radiology committee (BIRC)… Patients who were unmasked at radiographic progression and found to be in the placebo group could cross over, if eligible, to receive open-label cabozantinib. Patients in the cabozantinib group who had radiographic progression could also transition to open-label cabozantinib as long as they were deriving clinical benefit in the opinion of the investigator* [19]
If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Yes and No	ORR and safety data were recorded for these patients, and ORR was not assessed by BIRC after crossover (CS, B.2.3.2). There were some possible patient-reported outcomes for safety, e.g. pain, nausea, fatigue.[19]

Study question	Response (yes/no/not clear/NA) CS / EAG	How is the question addressed in the study? EAG
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No / No	-
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No (company-sponsored study) / No	The protocol published as a supplement with the principal manuscript reported all pre-specified outcomes.[19] It should be noted that the clinicaltrials.gov published protocol only reported the co-primary outcomes
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes/Yes Yes/Yes Yes/Yes	-

BIRC - blinded independent radiology committee; CS - company response; ITT - intention-to-treat; ORR - objective response rate; OS - overall survival; NA - not applicable

Table 8: Cochrane Risk of bias v.2.0: COSMIC-311 (based on data presented in Brose *et al.* 2021 and CS, Section B.2.3-2.5)

	Bias arising from the randomisation process: sequence generation, allocation concealment, balance between groups)	**Bias due to deviations from intended intervention (deviations with likely effect on outcomes)**	**Bias due to missing data (attrition)**	**Bias due to measurement of outcome (blinding of assessors, potential for differences between groups)**	**Bias in selection of reported results (pre-specified outcomes, potentially different measures)**	**Overall risk of bias**
Assessment	Low / Some concerns	High	Low	Low / Some concerns	Low	High
Details	*Randomisation was stratified by previous Lenvatinib treatment (yes vs no) and age (≤65 vs >65 years). The randomisation scheme used stratified permuted blocks of block size six and study treatment was centrally assigned through an interactive voice–web response system. Generation of the randomisation schedule was assigned to a clinical research organisation who maintained an unmasked team independent from the study. The live schedule, generated by the clinical research organisation, was uploaded to a secured server for the interactive response technology vendor who was responsible for interactive voice–web response services. Study personnel did not have access to the live schedule, the master list of blocks or block sizes, until authorised and documented unmasking (April 16, 2021)* (Brose 2021) Arms were balanced at baseline for known prognostic factors such as gender, but no information was provided on the prognostic factor of tumour volume/burden.[36, 37]	Open-label treatment with cabozantinib was permitted for eligible patients in either arm who progressed. Patients were analysed for OS and safety in the groups to which they had been randomised, thus compromising these data. The following patients in the placebo arm received post-progression, open-label cabozantinib: 19/62 (31%) at CCO1 (Brose 2021, Exelixis 2021a, Table 18) and 40/88 (45%) at CCO2 (Exelixis 2021b Section 3.2.1.1 and Exelixis 2021c, section 2.1 and Table 1). The following patients in the cabozantinib arm also received post-progression, open-label cabozantinib: 2/125 (1.6%) at CCO1 (Brose 2021) and 11/170 (6.5%) at CCO2 (Exelixis 2021c, Section 2.1 and Table 1)	ITT analyses were conducted	*The multiple primary endpoints were objective response rate in the first 100 randomly assigned patients (the objective response rate intention-to-treat [OITT] population) and progression-free survival in all randomly assigned patients (the intention-to-treat [ITT] population), both based on evaluations by BIRC* [19] ORR and safety data were recorded for unmasked patients, and ORR was assessed unblinded. There were some possible patient-reported outcomes for safety, e.g. pain, nausea, fatigue.[19]	The protocol published as a supplement with the principal manuscript reported all pre-specified outcomes.[19] It should be noted that the clinicaltrials.gov published protocol only reported the co-primary outcomes	As a result of the assessment of 'high' in one or more domains

BIRC - blinded independent radiology committee; CCO1 - clinical cut off 1 (19th August 2020); CCO2 - clinical cut off 2 (8th February 2021); CS – company's submission; CSR - Clinical Study Report; ITT - intention-to-treat; OITT - overall response rate intention-to-treat; ORR - overall response rate; OS - overall survival

The risk of bias due to deviations from the intended intervention was judged to be high due to the crossover of a large proportion of patients in the placebo arm within a relatively short period after commencement of the trial (31% at CCO1, 45% at CCO2, see Table 8), which affected key outcomes such as OS, despite analyses adjusting for crossover, and substantially reduced data on outcomes such as HRQoL, which was not recorded for crossover participants. As a result, the overall assessment of the EAG, following the Cochrane algorithm, is that the COSMIC-311 trial[19] is at high risk of bias.

4.2.3 Participant flow and analysis populations

Patients were randomised in a 2:1 ratio between February 2019 and February 2021. One hundred and seventy patients were assigned to the intervention arm (cabozantinib plus BSC) and 88 patients were assigned to the placebo control arm (BSC) (total N=258).[7] The trial reported two CCOs. The primary clinical cut-off, CCO1, was the 19th of August 2020, which related to 187 randomised subjects (125 cabozantinib, 62 placebo). This represents the ITT population. These results were reported in the principal publication[19] and the Clinical Study Report (CSR).[39]

After the 19th of August 2020 cut-off (CCO1), subjects continued to enrol in the study and receive blinded study treatment. The CS[7] reports that enrolment was stopped and the last subject was randomised on the 2nd of February 2021 because the study had demonstrated a significant improvement in PFS at the primary analysis (also referred to as the 'interim analysis', CS, Section B.2.4.2, Table 8). At this point, 258 randomised subjects (170 cabozantinib, 88 placebo) were enrolled in the trial. The second data cut-off date, CCO2, was the 8th of February 2021 for 'supportive analyses' (CS, Section B.2.4.1). This represents the Full ITT population. See Table 9 for the analysis populations.

Table 9: Analysis population for the COSMIC-311 trial

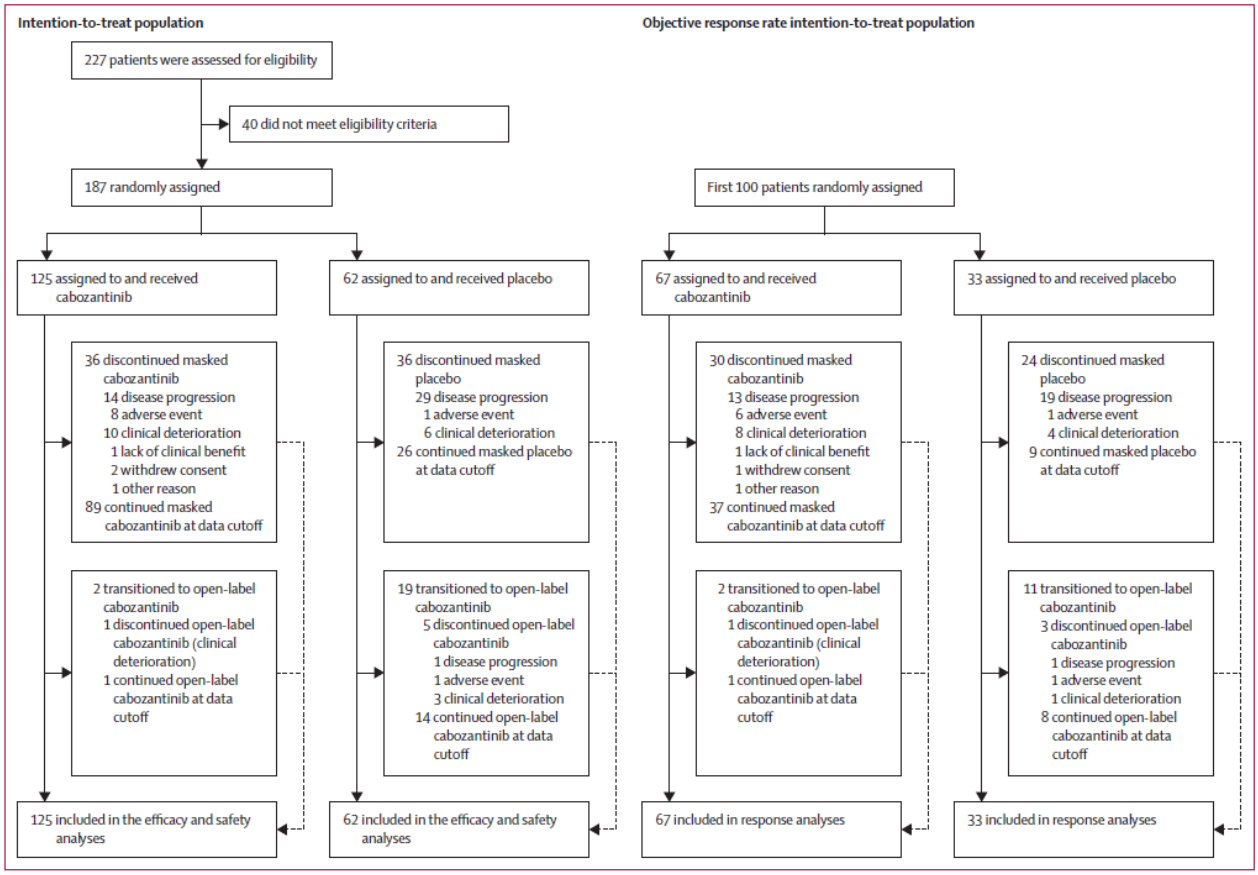
Analysis populations	**Number of patients**		
	Cabozantinib	**Placebo**	**Total**
CCO1 (19th August 2020)			
ITT population [19, 39] Safety [19, 39]	125	62	187
CCO2 (8th February 2021)			
Full ITT population [40] Safety[41]	170	88	258

CCO - clinical cut-off; ITT – intention-to-treat

It should be noted that patients were analysed in the groups to which they had been randomised despite the crossover phase, in which a proportion of unmasked placebo patients who had progressive disease received cabozantinib (see Figure 2). The following patients in the placebo arm received post-progression open-label cabozantinib: 19/62 (31%) at CCO1 (COSMIC-311[19] and XL184-311 CSR,[39] Table 18) and 40/88 (45%) at CCO2 (XL184-311 CSR Addendum 1,[40] Section 3.2.1.1 and XL184-311 CSR Addendum 2,[41] Section 2.1 and Table 1). The following patients in the cabozantinib arm also

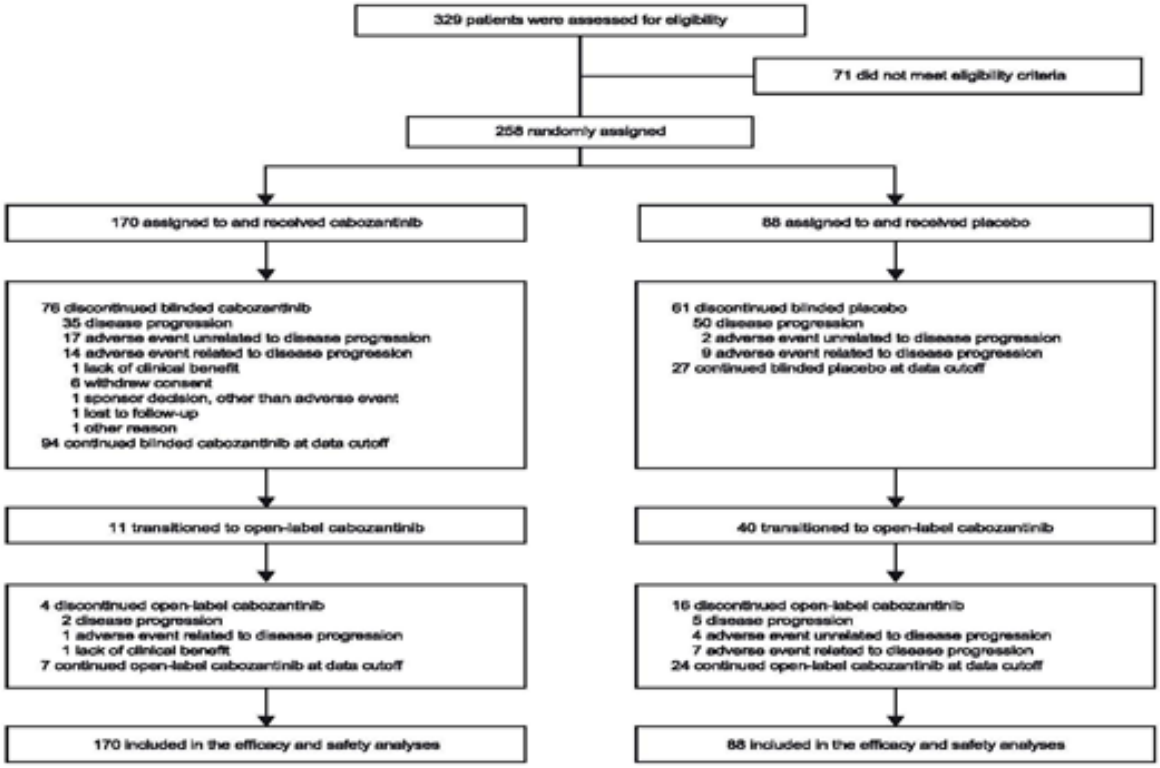
received post-progression open-label cabozantinib: 2/125 (1.6%) at CCO1 (Brose et al., 202119) and 11/170 (6.5%) at CCO2 (XL184-311 CSR Addendum 2,[41] Section 2.1 and Table 1). Full CONSORT diagrams of participant flow in COSMIC-311 for CCO1 and CCO2 are presented in Figure 3 and Figure 4, respectively.

Figure 3: Participant flow in COSMIC-311, CCO1 (reproduced from CS, Appendix D.1.2, Figure 2)



Source: Brose et al., 2021[19]

Figure 4: Participant flow in COSMIC-311, CCO2 (reproduced from CS, Appendix D.1.2, Figure 3)



4.2.4 Baseline characteristics in COSMIC-311

Participant characteristics in COSMIC-311[19] are presented in Table 10. In response to a question from the EAG regarding the company's statement that the COSMIC-311 trial participants are representative of DTC epidemiology, the company cited comparable data from the SELECT[42] and DECISION trials[43] on age only (see clarification response,[17] question A16). The company's response did not note any differences across other characteristics, for example, ECOG PS ≥2 status (COSMIC-311 across arms: 0%; SELECT 1.5-5%; DECISION 2.9-3.4 %) or the percentage of patients with bone metastases (COSMIC-311 39-50%; SELECT 37-40 %; DECISION 27-28%).

The reported characteristics were generally well-balanced between arms for both data-cuts CCO1 and CCO2. The EAG's clinical advisors stated that the COSMIC-311 trial population was generally consistent with the patients encountered in clinical practice in England. The EAG agrees that the two arms were balanced in terms of most known prognostic factors before the crossover phase, with the exception of tumour volume/burden, which was not controlled for.[36, 37] In response to a request for clarification from the EAG (question A18),[17] the company stated that they did not have data on this characteristic. Subject enrolment by country was reported in the CSR[39] (Section 10.1.4, Table 14) for

the CCO1 population and recorded four UK patients in the cabozantinib arm (3.2%) and three patients in the placebo arm (4.8%).

The CS[7] acknowledges that between 21% and 25% of participants in either arm of the trial at CCO2 and CCO1 had previously received both sorafenib and lenvatinib, which would not occur in practice in England given that both are only recommended by NICE as first-line treatment for this indication and there is currently no recommended second-line treatment (see Table 3). As a result, a proportion of the trial patients are not reflective of current practice in England.

Table 10: Characteristics of participants in COSMIC-311 across treatment groups (adapted from CS, Table 6, including data from Brose *et al.*, 2021 for CCO1)

COSMIC-311	**CCO2 (8th February 2021)**		**CCO1 (19th August 2020)**	
Baseline characteristic	**Cabozantinib**	**Placebo**	**Cabozantinib**	**Placebo**
Full ITT population	n=170	N=88	n=125	N=62
Age, median years (range)	65 (31-85)	66 (37-83)	65 (56-72)	66 (56-72)
[redacted]	[redacted]	[redacted]	63 (50)	33 (53)
Sex n (%)				
Male	83 (49)	39 (44)	57 (46)	28 (45)
Female	87 (51)	49 (56)	68 (54)	34 (55)
Geographical Region n (%)				
Europe	82 (48)	39 (44)	65 (52)	32 (52)
Asia	24 (14)	19 (22)	16 (13)	13 (21)
North America (USA and Canada)	15 (8.8)	12 (14)	13 (10)	9 (15)
Rest of the world	49 (29)	18 (20)	31 (25)	8 (13)
Race, n (%)				
White	121 (71)	59 (67)	90 (72)	41 (66)
Asian	29 (17)	20 (23)	20 (16)	14 (23)
Black or African American	2 (1.2)	2 (2.3)	1 (1)	2 (3)
Other / Not reported	18 (10.6)	7 (7.9)	14 (11)	5 (8)
ECOG PS, n (%)				
0 (normal activity, asymptomatic)	74 (44)	43 (49)	59 (47)	30 (48)
1 (fully ambulatory, symptomatic)	96 (56)	45 (51)	66 (53)	32 (52)
Smoking history, n (%)				
Current	[redacted]	[redacted]	NR	NR
Former				
Never				
Weight, median (range) (kg)			NR	NR
BMI, median (range) (kg/m2			NR	NR
Previous sorafenib or lenvatinib n (%)				
Sorafenib but no lenvatinib	61 (36)	33 (38)	46 (37)	23 (37)
Lenvatinib but no sorafenib	68 (40)	34 (39)	48 (38)	26 (42)
Sorafenib and lenvatinib	40 (23)	21 (24)	31 (25)	13 (21)
Other TKI therapy	1	0	NR	NR
Number of previous VEGFR				

COSMIC-311 Baseline characteristic	CCO2 (8th February 2021) Cabozantinib	 Placebo	CCO1 (19th August 2020) Cabozantinib	 Placebo
TKIs, n (%) 0 1 2	 1 (1) 126 (74) 43 (25)	 0 65 (74) 23 (26)	 0 91 (73) 34 (27)	 0 48 (77) 14 (23)
Histological subtype n (%)* Papillary Follicular	 96 (56) 78 (46)	 54 (61) 35 (43)	 67 (54) 62 (50)	 35 (56) 28 (45)
Metastatic lesions n (%) Bone Liver Lung Other	159 (94) 51 (30) 25 (15) 121 (71) 127 (75)	82 (93) 21 (24) 9 (10) 61 (69) 70 (80)	117 (94) 62 (50) 27 (22) 88 (70) 104 (83)	60 (97) 24 (39) 6 (10) 49 (79) 56 (90)
Number of prior PD-1/PD-L1 agents per subject for DTC, n (%) 0 1 ≥2 Median (range)	[redacted]	[redacted]	NR	NR
Median (range) time from progression on most recent prior non-radiation systemic anticancer regimen for DTC to randomisation, months	[redacted]	[redacted]	1.9 (1.0-4.0)†	1.9 (0.8-3.7)†

** Patients could be counted as having both papillary and follicular histological subtypes †IQR – values presented in parentheses reflect the inter-quartile range*
DTC - differentiated thyroid cancer; IQR - interquartile range; VEGFR TKI - vascular endothelial growth factor tyrosine kinase inhibitor; ITT - intention-to-treat; USA - United Sates of America; ECOG - Eastern Cooperative Oncology Group; PS - performance status; NR - not reported; PD-1 - programmed cell death protein 1; PD-L1 - programmed death ligand
Source: CS, B2.3.3, Table 6 and Brose et al., 2021

4.2.5 Study endpoints in COSMIC-311

The study endpoints with definitions are presented in Table 11. The primary efficacy analyses compared the results for the co-primary endpoints of PFS and ORR in subjects randomised to receive cabozantinib with those randomised to receive placebo. A number of 'additional' or 'secondary' endpoints were also assessed, including OS, HRQoL and safety. OS was defined in the CS[7] as being descriptive and non-inferential as it was not a controlled endpoint for the study; the primary purpose of the OS analyses was to evaluate the potential for detriment to survival with cabozantinib treatment (see CS, Section B.2.4.2). Pharmacokinetic (PK) measurements were also recorded for cabozantinib; these data are not reported here.

Table 11: Definitions of key outcome measures in COSMIC-311 (adapted from CS, Tables 7 and 8, and NCT03690388)

Outcome measure	Definition
Primary outcomes	
PFS	Time from randomisation to the date of radiographic disease progression as determined by the BIRC or death from any cause.
ORR	Time from randomisation to the BOR of confirmed CR or confirmed PR based on RECIST v1.1 as determined by the BIRC or investigator assessed for patients in the crossover phase (CS, Section B.2.3.2).
Key 'additional' outcomes	
OS	Time from randomisation to the date of death from any cause.
HRQoL	For HRQoL analyses, PROs were assessed using the EQ-5D-5L questionnaire, which is a 5-item, self-reported questionnaire comprised of 5 domains of health: mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, and anxiety and depression. Patients may indicate impairment in each domain according to five levels: no problems, slight problems, moderate problems, severe problems, and extreme problems.
DoR	Duration between the date of first documentation of response that is confirmed at least 28 days later to the earliest date of disease progression or death due to any cause.
Safety and tolerability	A TEAE was defined as any event with an onset date on or after the date of the first dose of study treatment or any ongoing event on the date of the first dose of study treatment that worsened in severity after the date of the first dose of study treatment. For brevity, "TEAE" is hereafter referred to as "AE." All AEs with an onset date through the end of the safety observation period were included in tabulations. AEs were considered study treatment-related if the Investigator determined that there was a possible relationship to the study treatment.

AE - adverse event; TEAE - treatment-emergent adverse event; BIRC - blinded independent radiology committee; BOR - best overall response; CR - complete response; EQ-5D-5L - EuroQol 5-Dimensions 5-level; HRQoL - health-related quality of life; ORR - objective response rate; OS - overall survival; PD - progressed disease; PR - partial response; RECIST - Response Evaluation Criteria in Solid Tumours; PFS - progression-free survival; SAE - serious adverse event; PRO - patient-reported outcome

4.3 Effectiveness of cabozantinib

The CS[7] presents efficacy results for both data cut-offs: the 19th of August 2020 (CCO1), which had a median follow-up of 6.2 months (interquartile range [IQR] 3.4-9.2) for the ITT population[19, 39] and the 8th of February 2021 (CCO2), with a median follow-up of 10.1 months for the Full ITT population (CS, Section B.2.12 and XL184-311 CSR Addendum 1[40]).

4.3.1 Progression-free survival (PFS) (co-primary endpoint)

PFS results are summarised in Table 12 (both data-cuts) and Figure 10 (CCO2 only). The COSMIC-311 trial reported significantly improved PFS for cabozantinib plus BSC compared to BSC alone at both data cut-offs (CCO1 and CCO2). At CCO1, median PFS was not reached for cabozantinib plus BSC, compared with 1.9 months for BSC alone (hazard ratio [HR] 0.22, 95% CI 0.13-0.36, p<0.001). At CCO2, median PFS was 11.0 months for cabozantinib plus BSC, compared with 1.9 months for BSC alone (HR 0.22, 95% confidence interval [CI] 0.15-0.32, p<0.001). However, a large proportion of

patients had censored data at CCO2 (64% in the cabozantinib arm and 22% in the placebo arm) and CCO1 (75% in the cabozantinib arm and 31% in the placebo arm).[7]

Table 12: PFS per BIRC, ITT population (reproduced from clarification response, Table 5)

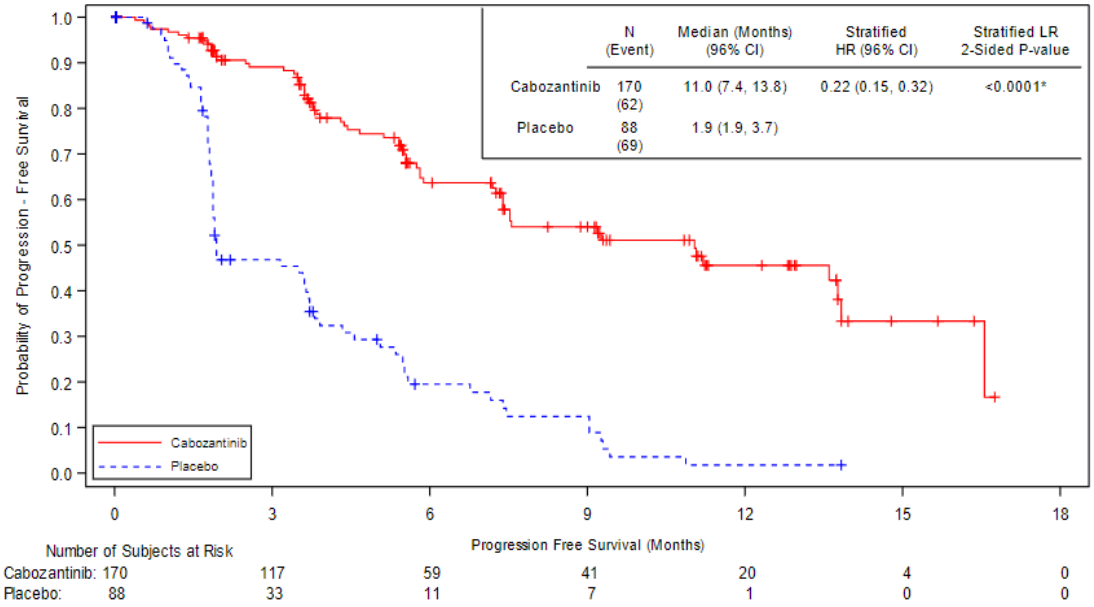
	CCO2 (Feb 2021) (N = 258)		CCO1 (Aug 2020) (N = 187)	
	Cabozantinib (N = 170)	**Placebo (N = 88)**	**Cabozantinib (N = 125)**	**Placebo (N = 62)**
Number (%) of subjects				
Censored	[redacted]	[redacted]	[redacted]	[redacted]
Receipt of local radiation to soft tissue for DTC	[redacted]	[redacted]	[redacted]	[redacted]
No post-baseline ATA[a]	[redacted]	[redacted]	[redacted]	[redacted]
No event by last ATA	[redacted]	[redacted]	[redacted]	[redacted]
2 or more missed ATA prior to event	[redacted]	[redacted]	[redacted]	[redacted]
Systemic NPACT	[redacted]	[redacted]	[redacted]	[redacted]
Event	62 (36)	69 (78)	31 (25)	43 (69)
Death	[redacted]	[redacted]	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]	[redacted]	[redacted]
Duration of PFS (months)				
Median (96% CI)	11.0 (7.4, 13.8)	1.9 (1.9, 3.7)	NE (5.7, NE)	1.9 (1.8, 3.6)
25th / 75th percentile[b]	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Observed *p*-value (stratified log-rank test)[c]	[redacted]		[redacted]	
HR (95% CI; stratified)[c,d]	[redacted]		[redacted]	
HR (96% CI; stratified)[c,d]	0.22 (0.15, 0.32)		0.22 (0.13, 0.36)	
Observed *p*-value (unstratified log-rank test)	[redacted]		[redacted]	
HR (95% CI; unstratified)[d]	[redacted]		[redacted]	
HR (96% CI; unstratified)[d]	[redacted]		[redacted]	
KM landmark estimates (% of subjects event-free) at:				
3 months	[redacted]	[redacted]	[redacted]	[redacted]
6 months	[redacted]	[redacted]	56.9	16.9
9 months	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]

ATA - adequate tumour assessment; BIRC - blinded independent radiology committee; CI - confidence interval; DTC - differentiated thyroid cancer; HR - hazard ratio; ITT - intention-to-treat; IxRS - interactive voice/web response system; KM - Kaplan-Meier; NPACT - non-protocol anticancer therapy; ORR - objective response rate; PD - progressive disease; PFS - progression-free survival

+ indicates a censored observation (see PFS censoring rules in XL184-311 CSR, Section 9.7.1.2.2)

a. In the Full ITT population, 11 cabozantinib and 8 placebo subjects were enrolled too close to the data cut-off date to have had a post-baseline tumour assessment. Four cabozantinib subjects withdrew from treatment before any post-baseline tumour assessment. Three cabozantinib subjects and 1 subject in the placebo arm died before their first post-baseline scan; b. Percentiles were based on KM estimates; c. Stratification factors: receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years). d. Estimated using the Cox proportional-hazard model (adjusted for stratification factors if applicable).

Figure 5: Kaplan–Meier plot of PFS per BIRC, CCO2 (reproduced from CS, Figure 5)



BIRC - blinded independent radiology committee; CI - confidence interval; HR - hazard ratio; ITT - intention-to-treat; IxRS - interactive voice/web response system; LR - log-rank test.
+ indicates value from censored observation

4.3.2 Objective response rate (ORR) (co-primary endpoint)

ORR results are summarised in Table 13. The COSMIC-311 trial reported a significantly improved objective response rate (ORR) for cabozantinib plus BSC compared with BSC alone at both data cut-offs (CCO1 and CCO2). At CCO1, the ORR was 9% (95% CI 4.5, 15.2) for cabozantinib plus BSC, compared with 0% (95% CI 0, 5.8) for BSC alone (p=0.017). At CCO2, the ORR was 11% (95% CI 6.9, 16.9) for cabozantinib plus BSC, compared with 0% (95% CI 0, 4.1) for BSC alone (p=0.0003). The secondary endpoint of duration of objective response (DoR) was also reported in the CS.[7] The Kaplan-Meier estimate of median (range) DoR per the Blinded Independent Radiology Committee (BIRC) was 10.2 months (1.87, 12.85 months) in the cabozantinib arm. The median (range) time from randomisation to the first objective response per BIRC was 3.6 (1.74, 7.52) months in the cabozantinib arm.[40]

Table 13: Objective response rate per BIRC, ITT population (reproduced from CS, Table 12)

	CC02 (N = 258)		CCO1 (N = 187)	
	Cabo (N = 170)	**Placebo (N = 88)**	**Cabo (N = 125)**	**Placebo (N = 62)**
Best overall response, n (%)[a]				
Confirmed CR	1 (0.6)	0	0	0
Confirmed PR	18 (10.6)	0	11 (9%)	0
SD	117 (68.8)	34 (38.6)	76 (61)	21 (34)
PD	11 (6.5)	42 (47.7)	8 (6)	31 (50)
No disease	1 (0.6)	0	1 (1)	0
Unable to evaluate	3 (1.8)	1 (1.1)	2 (2)	1 (2)
Missing	19 (11.2)	11 (12.5)	27 (22)	9 (15)
Objective response rate (CR+PR), n (%)	19 (11)	0	11 (9)	0
95% CI	6.9, 16.9	0.0, 4.1	4.5, 15.2	0.0, 5.8
Observed unstratified Fisher exact test *p*-value	0.0003		0.017	
Disease stabilisation rate (ORR+SD ≥ 16 weeks), n (%)	90 (52.9)	17 (19.3)	54 (43)	10 (16)
95% CI	[redacted]	[redacted]	34.4-52.5	8.0-27.7
DoR per BIRC (KM), median (range), months	10.2 (1.87+, 12.85+)	N/a	NR	N/a
Time to objective response per BIRC, median (range) time from randomisation, months[b]	3.581 (1.74, 7.52)	N/a	1.9 (1.8-3.6)	N/a

BIRC - blinded independent radiology committee; CI - confidence interval; CMH - Cochran Mantel-Haenszel; IxRS - interactive voice/web response system; RECIST - Response Evaluation Criteria in Solid Tumours; NR - not reached; N/a - not applicable

+ indicates censored observation

a) Best OR was based on RECIST 1.1 criteria and was calculated based on subjects in the OITT population. For CR or PR, confirmation of response must have occurred >28 days after the response was first observed.

b) Time to ORR (time from randomisation to the first subsequently confirmed CR or PR) is an arithmetic summary amongst those with an objective response and is defined as time from randomisation to the first CR or PR that is subsequently confirmed.

Source: XL184-311 CSR Addendum 1 (21st May 2021)[40] and Brose et al., 2021[19]

4.3.3 Overall survival (OS) (secondary endpoint)

OS results from COSMIC-311 are presented in Table 14 (both data-cuts) and Figure 6 (CCO2 only). The trial reported a non-significant trend in improved OS for cabozantinib plus BSC compared with BSC alone at both data cut-offs (CCO2 and CCO1): HR 0.54, 95% CI 0.27, 1.11, *p*=0.0879 (CCO1) and HR 0.76, 95% CI 0.45, 1.31, *p*=0.326 (CCO2). However, a large proportion of patients had censored data at CCO2 (78% in the cabozantinib arm and 76% in the placebo arm). The large proportion of placebo patients who, on progression, received open-label cabozantinib compromised the integrity of these data: 19/62 (31%) at CCO1 (COSMIC-311[19] and XL184-311 CSR ,[39] Table 18) and 40/88 (45%) at CCO2 (XL184-311 CSR Addendum 1,[40] Section 3.2.1.1).

Table 14: OS, ITT population (reproduced from clarification response, Table 6)

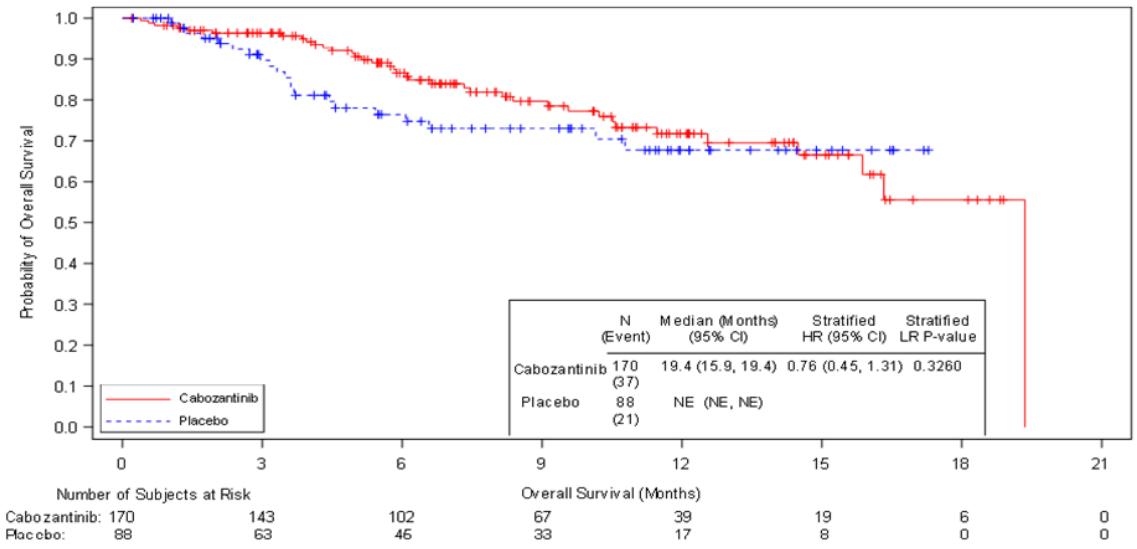
	CCO2		CCO1	
	Cabozantinib (N=170)	**Placebo (N=88)**	**Cabozantinib (N=125)**	**Placebo (N=62)**
Number of subjects (%)				
Censored	[redacted]	[redacted]	[redacted]	[redacted]
Alive	131 (77)	67 (76)	NR	NR
Death after data cut-off date	[redacted]	[redacted]	NR	NR
Death	37 (22)	21 (24)	17 (14)	14 (23)
Duration of overall survival (months)[a]				
Median (95% CI)	19.4 (15.9, NE)	NE (NE, NE)	NE (NE, NE)	NE (NE, NE)
25th percentile	[redacted]	[redacted]	[redacted]	[redacted]
75th percentile	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Observed *p*-value (stratified log-rank test)[b]	[redacted]		[redacted]	
HR (95% CI; stratified)[b,c]	0.76 (0.45, 1.31)		0.54 (0.27, 1.11)	
Observed *p*-value (unstratified log-rank test)	[redacted]		[redacted]	
HR (95% CI; unstratified)[c]	[redacted]		[redacted]	
KM landmark estimates (% of subjects event-free) at:				
3 months	[redacted]	[redacted]	[redacted]	[redacted]
6 months	[redacted]	[redacted]	84.8	73.4
9 months	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]

CI - confidence; HR - hazard ratio, ITT - intention-to-treat; LR - log-rank test, NE - not estimable; OS - overall survival; N/a - not applicable

+ Indicates censored observation; a Percentiles were based on K-M estimates; b Stratification factors based on IxRS were receipt of prior lenvatinib (yes vs no) and age at informed consent (≤ 65 years vs > 65 years); c Estimated using the Cox proportional-hazard model (adjusted for stratification factors if applicable). HR < 1 indicated OS in favour of cabozantinib; d In the Full ITT population and Primary Analysis subset, maximum duration of OS in the placebo arm was 17.28 months

Source: XL184-311 CSR (30th April 2020)[39] and XL184-311 CSR Addendum 1 (21st May 2021)[40] and Brose et al, 2021[19]

Figure 6: Kaplan-Meier plot of OS, CCO2 (reproduced from CS, Figure 9)



CI - confidence interval; HR - hazard ratio, ITT - intention-to-treat; LR - log-rank test, NE - not estimable; OS - overall survival

+ Indicates censored observation

In order to address potential confounding resulting from placebo group patients switching to receive open-label cabozantinib after progression, the company conducted a range of treatment switching analyses. The methods and results of these analyses are described in the context of the economic analysis in Section 5.2.4 of this EAG report.

Section B.2.6.4 of the CS[7] reports that there was no statistically significant difference between the two arms in terms of survival benefit, only a trend favouring cabozantinib, even after adjusting for crossover. The CS acknowledges that the results of these analyses were subject to bias and must be treated with caution.

4.3.4 Health-related quality of life

HRQoL of patients in COSMIC-311[19] was measured using the EuroQoL-5-Dimensions 5-Level (EQ-5D-5L) questionnaire and the EQ-5D Visual Analogue Scale (VAS). Patients completed the HRQoL assessments at baseline (before receiving the treatment or control), and every 4 weeks until week 25 and then every 8 weeks. EQ-5D-5L assessments were discontinued post progression and HRQoL results are only reported for CCO1 (following the clarification round, the company stated that additional HRQoL data are available for CCO2, although these have not been analysed or reported in the trial CSRs). A summary of the HRQoL results for both the EQ-5D-5L and the VAS scores from COSMIC-311 at CCO1 are presented in Table 15. The mean change from baseline in EQ-5D-5L and EQ-5D VAS is presented in Figure 7 and Figure 8, respectively. On all dimensions of the EQ-5D-5L, changes from baseline in patients in the cabozantinib and placebo arms did not show any statistically or clinically meaningful treatment difference (see CS,[7] Section B.2.6.5 for thresholds). The CS acknowledges that beyond week 33 there were fewer than 5 patients in the placebo arm, and any comparison for this period is therefore difficult to interpret. The CS states that these findings suggest that treatment of RAI refractory DTC patients with second- or third-line cabozantinib did not show a deterioration in HRQoL. However, this was only in the short-term (median follow-up for CCO1 was 6.2 months[19]).

Table 15: EQ-VAS and EQ-5D-5L scores - change from baseline, repeated measures analysis, ITT population, CCO1 (adapted from CS, Table 15)

	EQ-5D-5L Index		EQ-5D VAS	
Cabozantinib n (N=125)	[redacted]		[redacted]	
Cabozantinib LS mean (SE)				
Placebo n (N=62)				
Placebo LS mean (SE)				
Difference in mean change				
Pooled SD				
p-value				
Effect size				

EQ-5D-5L - EuroQol 5-Dimensions 5-Level; EQ-VAS - EuroQol visual analogue scale; SE - standard error; LS - least squares

Figure 7: Mean (SE) change from baseline EQ-5D-5L index score, CCO1, ITT population (reproduced from CS, Figure 12)

Post-BL - post-baseline; W - week
Source: XL184–311 CSR (30th April 2020)[39]

Figure 8: Mean (SE) change from baseline EQ-VAS score, CCO1, ITT population (reproduced from CS, Figure 13)

Post-BL - post-baseline; W - week
Source: XL184–311 CSR (30th April 2020)[39]

4.3.5 Subgroups

Pre-specified subgroup analyses were conducted for the primary efficacy outcome of PFS and the secondary outcome of OS. The patient sub-populations included: age; sex; race; location; ECOG PS; receipt of prior lenvatinib (yes vs. no); prior sorafenib (yes vs. no), and prior sorafenib and lenvatinib (yes vs. no); histology (papillary vs. follicular); and metastases: bone, important visceral, liver, lung metastases per investigator (yes vs. no).

The CS[7] reports that the overall PFS benefit was maintained across subgroups with reasonable sample sizes, including age and prior treatments (see Figure 9 for forest plots at CCO2). The CS acknowledges problems with these analyses, principally that the small sample sizes in some subgroups rendered the data difficult to interpret (CS, Section B.2.7). The CS also notes that, unlike for PFS, the HRs for OS did change across the overall population and subgroup populations in these analyses. The CS acknowledges that this, and the wide 95% CIs, were most likely due to the immaturity of the OS data (as few patients had had the event) and the problems created by treatment switching. Given the limitations of these analyses, the full results are not presented here, but are reported in the CS (Section B.2.7.2).

Figure 9: Forest plots of subgroup analyses for PFS per BIRC, CCO2, full ITT population, unstratified HRs (reproduced from CS, Figure 15)

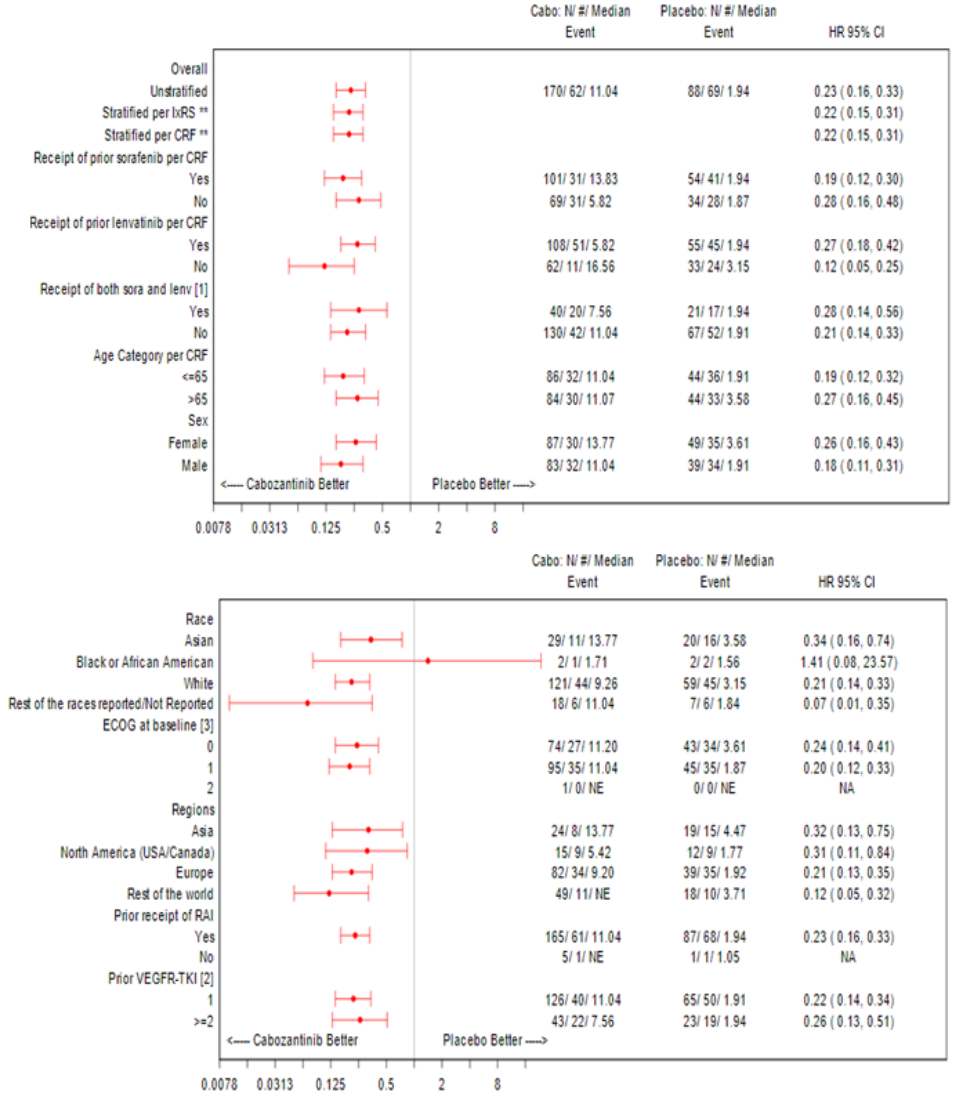
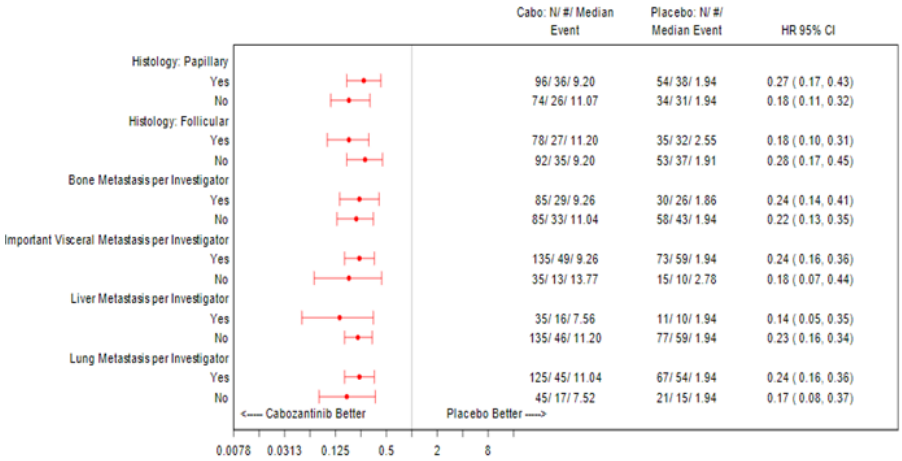


Figure 9 (cont'd): Forest plots of subgroup analyses for PFS per BIRC, CCO2, full ITT population, unstratified HRs (reproduced from CS, Figure 15)



BIRC - blinded independent radiology committee; Cabo - cabozantinib; CI - confidence interval; CRF - case report form; DTC - differentiated thyroid cancer; ECOG - Eastern Cooperative Oncology Group; PS - performance status; HR - hazard ratio; ITT - intention-to-treat; IxRS - interactive voice/web response system; NA - not applicable; NE - not estimable; PFS - progression-free survival; RAI - radioactive iodine; TKI - tyrosine kinase inhibitor; VEGFR - vascular endothelial growth factor receptor

4.4 Safety

4.4.1 Safety data reported for COSMIC-311

Section B.2.10 of the CS[7] reports safety data for the COSMIC-311 trial.[19] Data were reported for all randomised patients who had received at least one dose of treatment in either arm at both data cut-offs points (CCO1 and CCO2). However, it should be noted that there were sizeable proportions of patients in the placebo arm with progressive disease who crossed-over to receive open-label treatment with cabozantinib, which will have confounded safety data in the placebo arm: 31% at CCO1 (COSMIC-311[19] and XL184-311 CSR,[39] Table 18) and 45% at CCO2 (XL184-311 CSR Addendum 1,[40] Section 3.2.1.1 and XL184-311 CSR Addendum 2,[41] Section 2.1 and Table 1). This might explain some of the AEs reported in the control group despite the use of a placebo. A summary of the safety data from COSMIC-311 is presented in Table 16 and Table 17.

Treatment-related Grade 3 or 4 adverse events (AEs) were much higher in the cabozantinib arm than the placebo arm: respectively 47% vs. 6.5% (CCO1) and [redacted] vs.[redacted] (CCO2). Treatment-related serious adverse events (SAEs) were much higher in the cabozantinib arm than the placebo arm: respectively [redacted] vs. [redacted] (CCO1) and [redacted] vs. [redacted] (CCO2). AEs leading to dose reduction or interruption were much more common in the cabozantinib arm than the placebo arm: respectively [redacted] vs.[redacted] (CCO1) and [redacted] vs.[redacted] (CCO2).

Table 16: Overview of AEs, safety population, CCO1 and CCO2 (reproduced from CS, Table 17)

	CCO1		CCO2	
Parameters	**Cabozantinib (N=125) n (%)**	**Placebo (N=62) n (%)**	**Cabozantinib (N=170) n (%)**	**Placebo (N= 88) n (%)**
Any AE	117 (94)	52 (84)	166 (98)	75 (85)
Treatment-related AE	112 (90)	32 (52)	159 (94)	41 (47)
Grade 3 or 4 AE	71 (57)	16 (26)	106 (62)	25 (28)
Treatment-related Grade 3 or 4 AE	59 (47)	4 (6.5)	[redacted]	[redacted]
Grade 4 AE	7 (5.6)	2 (3.2)	11 (6.5)	2 (2.3)
Treatment-related Grade 4 AE	5 (4.0)	0	[redacted]	[redacted]
Grade 5 AE ≤ 30 days after last dose	[redacted]	[redacted]	[redacted]	[redacted]
Treatment-related Grade 5 AE ≤ 30 days after last dose	0	0	0	0
Treatment-related Grade 5 AE at any time	0	0	0	0
SAE	[redacted]	[redacted]	[redacted]	[redacted]
Treatment-related SAE	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose modification (reduction or interruption)	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose reduction	71 (57)	3 (4.8)	114 (67)	4 (4.5)
AE leading to dose interruption	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to treatment discontinuation (not related to disease under study)	6 (4.8)	0	15 (8.8)	0
Related to study treatment	[redacted]	[redacted]	[redacted]	[redacted]

**Patients are counted only once in each category but may be counted in multiple categories*
AE - adverse event; SAE - serious adverse event
For each treatment arm, the frequency and percentage of patients with AEs were tabulated by worst CTCAE grade for overall incidence by system organ class and preferred term or only by preferred term.

At CCO1, the most common (>20%) AEs in the cabozantinib arm were: [redacted] (XL184-311 CSR,[39] Table 54). The most common (>5%) Grade 3 or 4 AEs in the cabozantinib arm were: PPES (10% vs. 0% in the placebo arm); hypertension (9% vs. 3%); fatigue (8% vs. 0%); diarrhoea (7% vs. 0%); hypocalcaemia (7% vs. 2%).[19, 39] No Grade 3 or 4 AE occurred in the placebo arm at a frequency of >3%.

[redacted] At the CCO2 data cut-off, the most common (>5%) Grade 3 or 4 AEs in the cabozantinib

arm were: PPES (10% vs. 0% in the placebo arm); hypertension (12% vs. 2%); fatigue (9% vs. 0%); diarrhoea (8% vs. 0%); hypocalcaemia (8% vs. 2%) (XL184-311 CSR Addendum 2,[41] Table 8). Only dyspnoea occurred as a Grade 3 or 4 AE in the placebo arm at a frequency of >3% (3.4% vs. 1.8% in the cabozantinib arm) (XL184-311 CSR Addendum 2,[41] Table 8).

Table 17: Overview of most frequent AEs (>20% patients) in any arm or dataset, safety population, CCO1 and CCO2 (Brose *et al.* 2021, Exelixis 2021a and Exelixis 2021c)

	CCO1				CCO2			
Parameters	**Cabozantinib (N=125) n (%)**		**Placebo (N=62) n (%)**		**Cabozantinib (N=170) n (%)**		**Placebo (N=88) n (%)**	
	Any	**Grades 3/4**	**Any**	**Grades 3/4**	**Any**	**Grades 3/4**	**Any**	**Grades 3/4**
Diarrhoea	[redacted]	9 (7.2)	[redacted]	0	[redacted]	13 (7.6)	[redacted]	0
PPES	[redacted]	13 (10)	[redacted]	0	[redacted]	17 (10)	[redacted]	0
Hypertension	[redacted]	11 (8.8)	[redacted]	2 (3.2)	[redacted]	20 (12)	[redacted]	2 (2.3)
Fatigue	[redacted]	10 (8.0)	[redacted]	0	[redacted]	15 (8.8)	[redacted]	0
ALT increased	[redacted]	1 (0.8)	[redacted]	0	[redacted]	1 (0.6)	[redacted]	1 (1.1)
Nausea	[redacted]	4 (3.2)	[redacted]	0	[redacted]	4 (2.4)	[redacted]	0
AST increased	[redacted]	0	[redacted]	0	[redacted]	0	[redacted]	0
Decreased appetite	[redacted]	4 (3.2)	[redacted]	0	[redacted]	5 (2.9)	[redacted]	0
Hypocalcaemia	[redacted]	9 (7.2)	[redacted]	1 (1.6)	[redacted]	13 (7.6)	[redacted]	2 (2.3)
Weight decreased	[redacted]	1 (0.8)	[redacted]	0	[redacted]	4 (2.4)	[redacted]	0

Source: CCO1: CSR and Brose 2021; CCO2: CSR Addendum 2

4.4.2 Safety summary

Some AEs are frequent in patients receiving cabozantinib plus BSC ([redacted]), principally diarrhoea, PPES, hypertension, fatigue and hypocalcaemia. These AEs also occur in patients receiving cabozantinib plus BSC at Grades 3/4 AEs, but at relatively low frequencies (<12%).

4.5 Additional study of cabozantinib

A second trial that satisfied the inclusion criteria for the SLR was identified in the CS[7] but was excluded because some patients experienced dose escalation of cabozantinib above the licensed dose (80mg rather than 60mg) (NCT01811212, reported in Cabanillas *et al.*, 2017[34]) (CS, Section B.2.8). This exclusion was queried by the EAG and the company subsequently agreed to the inclusion of this trial and its data (see clarification response,[17] question A9). Details of this trial are presented below.

This was a single-arm, multicentre study of adult patients with RAI-refractory DTC (N=25) who had received at least one line of prior VEGFR-targeted therapy. It was conducted in six centres in the USA. Cabozantinib was administered orally at a starting dose of 60mg daily in 28-day cycles. Patients who tolerated cabozantinib with no Grade ≥2 treatment-related AEs could have their dose increased to 80mg

daily. Those patients experiencing Grade ≥2 treatment-related AEs had their dose reduced to 40mg daily (and again to 20mg daily, if necessary). Treatment was continued until disease progression, unacceptable toxicity or withdrawal of consent. Seven patients (28%) were treated at 60mg/day of cabozantinib, whereas four (16%) had a dose escalation to 80mg/day and 14 (56%) had a dose reduction to 40mg/day (n=6; 24%) or to 20mg/day (n=8; 32%).

A comparison of the baseline characteristics of patients enrolled in NCT01811212 and the cabozantinib arm of COSMIC-311 is presented in Table 18. Patient characteristics, where comparison was possible, were generally similar between NCT01811212 and COSMIC-311 except for sex (64% in NCT01811212 vs. 51% male in COSMIC-311, respectively), histology, e.g., follicular subtype (16% vs. 46%) and the proportions of patients with the following metastases: bone (84% vs 30%) and liver (36% vs. 15%).

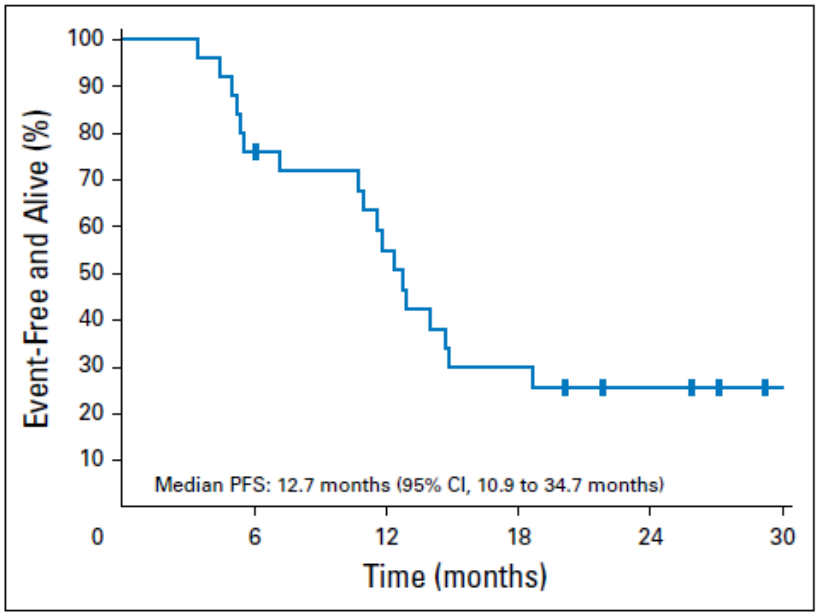
Table 18: Characteristics of participants in NCT01811212 and the cabozantinib arm of COSMIC-311

	NCT01811212	**COSMIC-311**
	Cabozantinib	**Cabozantinib arm (CCO2)**
Full ITT population	n=25	N=170
Age, median years (range)	64 (41-81)	65 (31-85)
Sex n (%)		
Male	16 (64)	83 (49)
Female	9 (36)	87 (51)
Geographical Region n (%)		
Europe	-	82 (48)
Asia	-	24 (14)
North America (USA and Canada)	25 (100)	15 (8.8)
Rest of the world	-	49 (29)
Race, n (%)		
White	21 (84)	121 (71)
Asian	3 (12)	29 (17)
Black or African American	1 (4)	2 (1.2)
Number of previous vascular endothelial growth factor receptor tyrosine kinase inhibitors n (%)		
0	-	1 (1)
1	21 (84)	126 (74)
2	4 (16)	43 (25)
Histological subtype n (%) 1		
Papillary	9 (36)	96 (56)
Follicular	4 (16)	78 (46)
Metastatic lesions n (%)	-	159 (94)
Bone	21 (84)	51 (30)
Liver	9 (36)	25 (15)
Lung	21 (84)	121 (71)
Other	-	127 (75)

DTC - differentiated thyroid cancer; IQR - interquartile range; ITT - intention-to-treat; USA - United Sates of America; ECOG - Eastern Cooperative Oncology Group; PS - performance status; NR - not reported; PD-1 - programmed cell death protein 1; PD-L1 - programmed death ligand
Source: CS, B2.3.3, Table 6 and Brose et al., 2021.

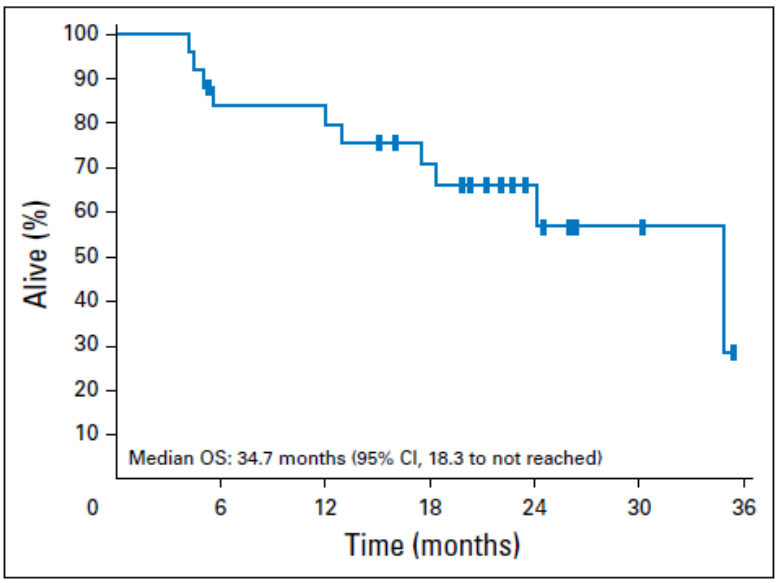
The median duration of follow-up was 22.8 months (95% CI, 21.2, 30.2 months), compared with 10.1 months for COSMIC-311 at CCO2 (Exelixis 2021b). Median PFS was 12.7 months (95% CI, 10.9 to 34.7 months) (see Figure 10), and the estimated PFS rate was 55% (95% CI, 38% to 79%) at 12 months and 25% (95% CI, 13% to 50%) at 24 months. Median OS was 34.7 months (95% CI, 18.3 months to not reached) and the estimated OS rate was 80% (95% CI, 65% to 97%) at 12 months and 66% (95% CI, 49% to 88%) at 24 months (see Figure 11).

Figure 10: Kaplan-Meier plot of PFS, Study NCT01811212



PFS - progression-free survival; CI - confidence interval

Figure 11: Kaplan-Meier plot of OS, Study NCT01811212



OS - overall survival; CI - confidence interval

Cabozantinib-related Grade 3 AEs experienced by >5% of patients in this trial were: hypophosphatemia (16%), fatigue, weight loss, neutropenia and lipase or amylase elevation (12%); and diarrhoea, PPES, hyponatremia and hypokalaemia (8%).[34] The following SAEs were also noted: Grade 1 thrombotic thrombocytopenic purpura (n=1), Grade 2 deep venous thrombosis (n=1), Grade 4 perianal hidradenitis suppurativa (n=1), and Grade 3 AEs (n=6) comprising left ventricular systolic dysfunction, asymptomatic increased lipase, osteonecrosis of the jaw, decubitus ulcer, pneumonia, and meningitis (one of each event).[34] It was also recorded that there was one death *"possibly attributable to cabozantinib."*[34] Grade 1 and 2 (combined) AEs with a frequency of >50% were: fatigue, anorexia, oral mucositis, nausea, diarrhoea, PPES, liver transaminase elevation, and hypomagnesemia. The high frequencies of diarrhoea, PPES, fatigue, nausea and liver transaminase elevation were also noted in the COSMIC-311 trial.[41]

4.6 Ongoing studies

The CS[7] states that there were no relevant ongoing studies of cabozantinib in this population that are likely to report in the next 12 months. This is correct for cabozantinib as monotherapy. Response and safety outcomes are to be reported in a single-arm, Phase II study of cabozantinib in combination with nivolumab and ipilimumab in RAI-refractory adult DTC patients who have progressed on one line of VEGFR-targeted therapy (including but not limited to sorafenib, sunitinib, vandetanib, pazopanib, or lenvatinib, etc.). However, details of the dose and regimen are not reported in the published protocol. This trial currently has a primary completion date of January 2023 (NCT03914300).

In addition, the company's clarification response[17] (question A20) notes that no further data-cuts of COSMIC-311 are planned.

4.7 Meta-analysis

Section B.2.8 of the CS[7] states reports that no meta-analysis was conducted. This was because two additional Phase II trials of cabozantinib in adult patients with RAI-refractory DTC were identified in the CS, but both were deemed not to be relevant to the decision problem. The first Phase II trial was considered not to be relevant because it evaluated cabozantinib as first-line rather than second-line treatment (NCT02041260), and the second trial[34] (NCT01811212) was considered not to be relevant because it permitted cabozantinib dose escalation to 80mg (rather than the licensed 60mg dose as per the decision problem). However, as noted above, in response to the EAG's clarification request, the company considered this latter trial to be potentially relevant and briefly summarised the study and its findings in their clarification response[17] (question A9). The company did not conduct a meta-analysis with this trial and the COSMIC-311 trial.[19]

4.8 Indirect treatment comparisons

The CS[7] reports that no ITC was conducted. This was because trials of alternative therapies for adult patients with RAI-refractory DTC only considered treatments that were not recommended in the UK as second-line therapies (e.g., lenvatinib and sorafenib), did not present sufficient evidence for subsets of participants who received the therapy at second-line (<5% of patients in the sorafenib DECISION trial and with no subgroup data[43]; and ≤25.3% in any arm in the lenvatinib SELECT trial, but only with subgroup data for PFS,[42] or only considered a therapy that was licensed for a specific subgroup of patients with RET-mutation (selpercatinib) (CS, Section B.2.9). In the absence of such data, the most relevant comparator was deemed to be BSC, and the COSMIC-311 trial provides direct data on that comparison.

4.9 Additional work on clinical effectiveness undertaken by the EAG

The EAG did not undertake any additional work relating to the clinical effectiveness of cabozantinib.

4.10 Conclusions of the clinical effectiveness section

The pivotal trial of cabozantinib, COSMIC-311,[19] was an international, multicentre, randomised, placebo-controlled, blinded, Phase III trial. The trial had two CCOs: the primary clinical cut-off date was the 19th of August 2020 (number of patients: 125 cabozantinib, 62 placebo) (CCO1); and, after further enrolment, the second, 'supportive analyses' data cut-off date was the 8th of February 2021 (170 cabozantinib, 88 placebo) (CCO2). The study was assessed by the EAG as being at high risk of bias on account of the deviation from the pre-specified interventions. This was due to the sizeable proportions of patients in the placebo arm with progressive disease who crossed-over to receive open-label treatment with cabozantinib therapy (31% at CCO1 and 45% at CCO2), which confounded the outcomes of OS and safety. COSMIC-311 was a medium-sized trial with 258 subjects across two arms at CCO2, but with a short length of follow-up (median 10.1 months at the latest data cut-off, CCO2, and 6.2 months at the primary data cut-off, CCO1). The CS[7] reports that there are no plans to conduct further data-cuts beyond CCO2 (clarification response,[17] question A20).

Cabozantinib demonstrated significant efficacy compared with placebo in terms of PFS and ORR at both data cut-offs. However, the outcomes of OS and safety were confounded by the short time to progression for patients in the placebo arm (median 1.9 months at both CCO1 and CCO2, see Table 12) combined with the high levels of censoring and the crossover design. The CS[7] reports that there was no significant difference between the two arms in terms of OS, only a trend favouring cabozantinib, even after adjusting for crossover.

There were high rates of treatment-related AEs and SAEs in the cabozantinib arm compared with the placebo arm, as well as dose modifications due to AEs. A number of AEs related to cabozantinib

treatment were frequent: diarrhoea, PPES, hypertension, fatigue, hypocalcaemia, [redacted], and decreased appetite. Some of these AEs were also the most frequent at Grade 3 or higher, but were not common (<12%). HRQoL was only assessed by EQ-5D-5L in patients who had not progressed / up to the point of progression (to prevent confounding due to crossover) and no significant or clinically important difference between cabozantinib and placebo was found for patients who had not progressed up to 33 weeks (there were only five or fewer patients in the placebo arm after this point, preventing meaningful comparisons from being made). The CS[7] interprets this finding as a lack of detriment to HRQoL from cabozantinib-related AEs.

No meta-analysis was conducted despite the presence of a single-arm trial that satisfied the SLR inclusion criteria, and no ITC was undertaken because of the absence of comparable trials of second-line therapy in the relevant population, and the availability of direct evidence from a single Phase III RCT comparing cabozantinib at the licensed dose of 60mg daily[20] with the comparator listed in the final NICE scope[18] (BSC).

5. COST EFFECTIVENESS

This chapter presents a summary and critique of the company's health economic analyses of cabozantinib for the treatment of patients with locally advanced or metastatic DTC, whose disease is refractory to, or who are unsuitable for RAI, and whose disease has progressed during or after prior systemic therapy. Section 5.1 describes and critiques the company's SLR of existing economic analyses of treatments for RAI-refractory DTC. Sections 5.2 describes the company's economic model and summarises the company's results. Sections 5.3 presents the EAG's critical appraisal of the company's original economic model. Section 5.4 briefly summarises and critiques the company's updated model provided following the clarification round. Section 5.5 presents the methods and results of the EAG's exploratory analyses. Section 5.6 discusses the key issues around the company's economic analysis.

5.1 Critique of the company's review of existing economic analyses

5.1.1 Summary and critique of the company's searches

The company undertook an SLR of existing economic studies of treatments for RAI-refractory locally advanced or metastatic DTC. CS Appendix G[7] reports the searches conducted to inform the company's review of existing economic studies, as well as those used to inform the company's reviews of HRQoL studies and cost and resource use studies (reported in CS Appendices G, H and I respectively).

These searches, which cover an appropriate selection of databases and conference proceedings, were conducted shortly after those for the clinical SLR (14th October 2021), so as with the clinical SLR, there is a considerable gap between the process of evidence identification and the date of submission to NICE and the EAG (20th September 2022).

The company's searches are well-designed, combining the same population terms as the clinical searches with filters based on the work of the SIGN, the Canadian Agency for Drugs and Technologies in Health (CADTH) and the York Health Economics Consortium (YHEC). As with the clinical review, the main searches were conducted as a multi-file search (this time across MEDLINE, Embase, Econlit and PsycINFO simultaneously) meaning that it was not possible for the EAG to replicate the searches exactly as they were conducted or to assess the impact of this searching approach on study retrieval.

Despite potential concerns about the lack of formal update searches and the difficulties of reproducing the company's multi-file search approach, the EAG is broadly satisfied that the company has made a reasonable attempt to identify all the relevant evidence up to the point at which the searches were conducted.

5.1.2 Summary and critique of company's review of existing economic evaluations

The company's review included six studies of treatments for patients with RAI-refractory, locally advanced or metastatic DTC (see Table 19). Only one study was available as full text;[44] the other five included studies were available as abstracts only. Five of the studies were cost-utility analyses which reported outcomes in terms of the incremental cost per quality-adjusted life year (QALY) gained; the sixth study, Carlson *et al.*,[45] does not report costs or incremental cost-effectiveness ratios (ICERs) and therefore should have been excluded according to the eligibility criteria for the review (see CS Appendix G,[25] Table 37). The included studies assessed a range of interventions and comparators including sorafenib, lenvatinib, larotrectinib and placebo/BSC. None of the included studies evaluated cabozantinib for RAI-refractory DTC. Three of the included studies adopted a state transition modelling approach whilst the remaining three studies were partitioned survival models. As most of the studies were available only in abstract form, few details are available regarding the models and their assumptions. The EAG considers that none of the identified studies were sufficient to address the decision problem for this appraisal and that a *de novo* model was required.

Table 19: Summary of studies included in company's review of economic analyses

Study	Publication type	Population	Interventions/ comparators	Outcome	Setting	Model type
Erdal *et al.* (2015)[46]	Abstract	RAI-refractory locally advanced/ metastatic DTC	• Sorafenib • BSC	Incremental cost per QALY gained	Turkey	PartSA
Tremblay *et al.* (2016)[47]	Abstract	RAI-refractory DTC	• Lenvatinib • Sorafenib	Incremental cost per QALY gained	US	PartSA
Huang *et al.* (2016)[48]	Abstract	RAI-refractory DTC	• Lenvatinib • Sorafenib • Placebo	Incremental cost per QALY gained	US	STM
Wilson *et al.* (2017)[44]	Full text	Progressed RAI-refractory DTC	• Lenvatinib • Sorafenib • Placebo	Incremental cost per QALY gained	US	STM
Carrasquilla-Sotomayor *et al.* (2017)[49]	Abstract	RAI-refractory DTC	• Sorafenib • BSC	Incremental cost per QALY gained	Columbia	STM
Carlson *et al.* (2021)[45]	Abstract	NTRK positive RAI-refractory DTC	• Larotrectinib • Sorafenib • Lenvatinib	Incremental QALYs gained	US	PartSA

RAI - radioactive iodine; DTC - differentiated thyroid cancer; BSC - best supportive care; US - United States; QALY - quality-adjusted life year; STM - state transition model; PartSA - partitioned survival model; NTRK - neurotrophic tyrosine receptor kinase

5.2 Description of the company's original economic analysis

This section describes the company's original submitted economic model, as described in the CS.[7] Following the clarification round, the company submitted a revised economic analysis which includes a number of amendments to the model assumptions and parameters. This revised model is described and critiqued separately in Section 5.4.

5.2.1 Scope of the company's economic analysis

As part of their submission to NICE,[7] the company submitted an executable health economic model programmed in Microsoft Excel.® The scope of the company's economic analysis is summarised in Table 20.

Table 20: Scope of the company's economic analysis

Population	Adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed during or after prior systemic therapy
Time horizon	35 years (lifetime)
Intervention	Cabozantinib 60mg QD (administered orally)
Comparator	BSC
Type of economic analysis	Cost-utility analysis
Outcome	Incremental cost per QALY gained
Perspective	NHS and PSS
Discount rate	3.5% per annum
Price year	2020/21 (except for drug costs which reflect current prices)

DTC - differentiated thyroid cancer; RAI - radioactive iodine; mg - milligram; QD - once a day; QALY - quality-adjusted life year; NHS - National Health Service; PSS - Personal Social Services; BSC - best supportive care

The company's economic model assesses the cost-effectiveness of cabozantinib versus BSC for the treatment of adult patients with locally advanced or metastatic DTC who are refractory to or not eligible to receive RAI and who have progressed during or after prior systemic therapy. Cost-effectiveness is assessed in terms of the incremental cost per QALY gained from the perspective of the NHS and Personal Social Services (PSS) over a 35-year horizon. Unit costs are valued at 2020/21 prices, except for drug acquisition costs which are valued at current prices. Health outcomes and costs are discounted at a rate of 3.5% per annum.

Population

The company's economic analysis reflects the full ITT population of the COSMIC-311 trial.[19] As noted in Section 4.2.4, in COSMIC-311, approximately 76% of patients had previously received either sorafenib or lenvatinib, whilst the remaining 24% of patients had received both of these TKIs. At model entry, patients are assumed to be 65 years of age and 53% of patients are assumed to be female.

The EAG's clinical advisors commented that the trial population is broadly representative of the DTC patient population who would be offered cabozantinib if it was available on the NHS. Both advisors

stated that they typically see similar proportions of men and women and one advisor mentioned that patients who would be treated in the NHS would likely be slightly younger than the trial population. The clinical advisors commented that patients treated in the NHS would not receive both lenvatinib and sorafenib. One advisor also commented that they would expect a pure second-line population to have *"very slightly"* better outcomes compared with the COSMIC-311 population.

Intervention

The intervention included in the company's economic analysis is cabozantinib, administered orally at a dose of 60mg once daily. This is in line with the final NICE scope[18] and the EMA/MHRA marketing authorisation for cabozantinib for the DTC indication.[20] The SmPC for cabozantinib[20] (page 3) states that *"Treatment should continue until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs."* However, the company's base case model includes a structural constraint which forces time to treatment discontinuation (TTD) to be less than or equal to PFS; hence, the model implicitly assumes that patients will discontinue treatment with cabozantinib at the point of disease progression. The company's clarification response[17] (question A15) states that the company is seeking a positive NICE recommendation for the use of cabozantinib in line with the SmPC, which permits treatment beyond progression if the patient is still deriving benefit from it and if they are not experiencing toxicity. The EAG notes that at the CCO2 data cut-off (8th February 2021), 11 of 170 patients (6.5%) randomised to the cabozantinib arm of COSMIC-311 had received open-label cabozantinib after disease progression.[41] The company's model does not include any adjustment of the OS data for the cabozantinib group of COSMIC-311 to account for the potential additional benefit of continued cabozantinib treatment received after disease progression in the trial, nor does it include the costs of cabozantinib given after progression in the trial. However, the proportion of patients who received open-label cabozantinib is small and the EAG's clinical advisors agreed that the impact of any potential confounding on OS is likely to be minor. The model assumes that patients do not receive any further active anticancer therapy after progressing on cabozantinib (i.e., they receive BSC alone). The model includes a PAS discount for cabozantinib of [redacted].

Comparators

The company's economic analysis includes a single comparator – BSC. This is consistent with the final NICE scope.[18] Outcomes for BSC are modelled using data from the placebo arm of COSMIC-311,[19] including the statistical adjustment of OS to account for confounding resulting from placebo patients switching to receive cabozantinib after disease progression (40 of 88 patients [45.5%] at CCO2). The company's model includes BSC costs associated with: tests (urea and electrolytes, haematology/coagulation, calcium and magnesium, liver function and thyroid function); scans (electrocardiograms [ECGs] and computerised tomography [CT]); clinical consultations (consultant-led and nurse-led) and end-of-life care (see Section 5.2.4). These costs are applied to the health states

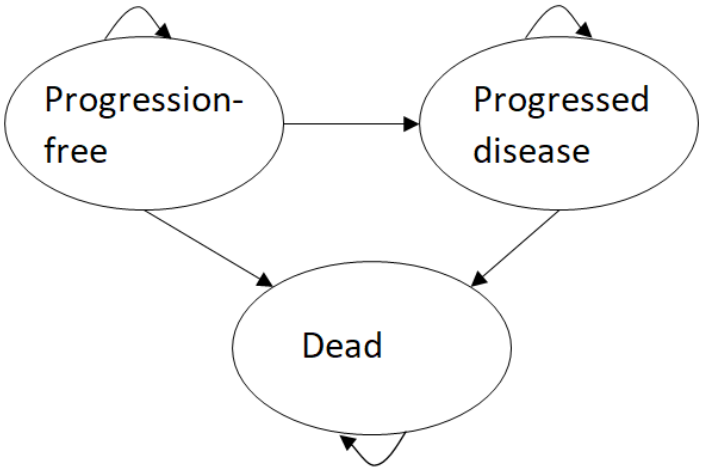
in both treatment groups. The costs of TSH suppression and other concomitant therapies are not included.

[redacted][21] The EAG's clinical advisors also stated that they most commonly use lenvatinib rather than sorafenib as first-line therapy. The EAG's clinical advisors stated that they discontinue TKI treatment at the point of disease progression, although one advisor commented that they might offer continued lenvatinib treatment if progression is limited to a single site. They agreed that some other clinicians offer continued treatment if the patient is still obtaining clinical benefit. Continued lenvatinib given after progression is not considered as a comparator in the company's economic analysis; this issue is discussed further in Section 5.3.5.

5.2.2 Model structure and logic

The company's economic model adopts a partitioned survival approach, including three health states: (i) progression-free; (ii) progressed disease, and (iii) dead (see Figure 12).

Figure 12: Company's model structure



The model logic operates as follows. Patients enter the model in the progression-free state and receive treatment with either cabozantinib (plus BSC) or BSC alone. At any time *t*, health state occupancy is determined by the cumulative probabilities of OS and PFS, whereby: the probability of being alive and progression-free is given by the cumulative probability of PFS; the probability of being alive following disease progression is calculated as the cumulative probability of OS minus the cumulative probability of PFS, and the probability of being dead is calculated as one minus the cumulative probability of OS. The company's model includes half-cycle correction, although this is subject to an error (see Section

5.3.5). Patients in the cabozantinib group are assumed not to receive treatment after disease progression, based on the TTD function which has been capped by PFS. No further active anticancer treatments are assumed to be given after disease progression in the cabozantinib group, or to any patient in either alive health state in the BSC group.

The cumulative probabilities of OS and PFS for patients receiving cabozantinib and BSC are modelled using parametric survival models fitted to time-to-event data from the COSMIC-311 trial.[19] The economic model applies a structural constraint which ensures that the cumulative probability of survival in the target RAI-refractory DTC population cannot be higher than that in the age- and sex-matched general population. However, this aspect of the model is subject to errors (see Section 5.3.5). The model also applies a constraint which ensures that the cumulative probability of PFS cannot be higher than the cumulative probability of OS. The model applies a further structural constraint which forces all BSC-treated patients who are still alive at 5 years to move to the dead state at this timepoint.

HRQoL is assumed to be determined by the presence/absence of disease progression, with the same utility values applied in each treatment group. Utility values are adjusted for increasing age. The model also includes short-term QALY losses associated with Grade 3/4 treatment-emergent adverse events (TEAEs) which occurred in ≥5% of either arm in COSMIC-311.[19] TEAEs are assumed to have a negative HRQoL impact for a duration of one month.

The model includes costs associated with: (i) drug acquisition; (ii) prescribing and dispensing of cabozantinib; (iii) health state management (scans, tests and clinic visits); (iv) the management of AEs and (v) end-of-life care costs. Drug acquisition costs for cabozantinib are modelled as a function of the TTD distribution (constrained by PFS), the treatment schedule,[20] treatment compliance in COSMIC-311[19] and the PAS-discounted price. Health state costs are applied in each model cycle. End-of-life care costs are applied once-only at the point of death.

The incremental health gains, costs and cost-effectiveness for cabozantinib versus BSC are estimated over a 35-year time horizon using monthly cycles. No economic subgroup analyses are presented in the CS.[7]

5.2.3 Key assumptions employed in the company's model

The company's economic model employs the following key assumptions:

- The modelled population is 65 years of age at model entry.[19]
- BSC is the sole comparator for cabozantinib.
- PFS is modelled using independent Weibull distributions fitted to the observed PFS data from COSMIC-311 for cabozantinib and placebo.

- OS is modelled using independent exponential distributions fitted to the observed OS data for cabozantinib and the RPSFT-adjusted data for placebo (including re-censoring).
- Time on treatment is modelled using an exponential distribution fitted to data on TTD, linked to the date of the last known dose (see clarification response,[17] question A15). This implies a stopping rule whereby cabozantinib is discontinued in all patients who are still receiving cabozantinib at the point of disease progression. Post-progression cabozantinib use in COSMIC-311[19] is assumed to not have impacted on OS and the costs associated with post-progression cabozantinib use in the trial are not included in the economic model.
- The model includes three structural constraints: (i) the cumulative probability of OS with DTC cannot be higher than that in the age- and sex-matched general population; (ii) the cumulative probability of PFS cannot be higher than the cumulative probability of OS, and (iii) all BSC-treated patients who are still alive at 5 years will die at this timepoint. Given the use of a partitioned survival approach, PFS and OS are otherwise structurally unrelated.
- HRQoL is dependent on the presence/absence of disease progression. The same utility values are applied to the health states in each treatment group. The utility value for the progression-free state is higher than that for progressed disease state. Utility values are age-adjusted but are not capped by general population EQ-5D values.
- AEs result in QALY losses and additional costs. AEs are assumed to be resolved by the end of the first 1-month model cycle.
- Prior to progression, disease management costs are slightly higher for patients receiving cabozantinib compared with those receiving BSC alone. For patients who have progressed, the same disease management cost is applied in both treatment groups.

5.2.4 Evidence used to inform the company's model parameters

Table 21 summarises the evidence sources used to inform the model parameter values. The evidence sources and the derivation of the parameter values are described in detail in the subsequent sections.

Table 21: Summary of evidence used to inform the company's original base case model

Parameter / group	Cabozantinib	BSC
Patient characteristics (age and sex)	COSMIC-311[19]	
PFS	Weibull model fitted to cabozantinib group PFS data from COSMIC-311[19]	Weibull model fitted to BSC group PFS data from COSMIC-311[19]
OS	Exponential model fitted to cabozantinib group OS data from COSMIC-311[19]	Exponential model fitted to BSC group (RPSFT-adjusted) OS data from COSMIC-311.[19] 5-year death assumption based on input obtained from company's 2022 advisory board meeting.[21]
TTD	Exponential model fitted to TTD in COSMIC-311,[19] capped by PFS model	N/a
General population mortality	Arbitrary numbers used in executable model (see Section 5.3.5)	
Health state utility values	Fordham *et al.* general population TTO study[22] (adjusted values from multivariable regression model).	
TEAE frequencies	Grade 3/4 TEAEs arising in ≥5% of patients in either treatment group in COSMIC-311[19]	
TEAEs disutilities	Fordham *et al.* TTO study[22] and the AXIS trial[23] (axitinib for RCC)	
TEAE duration	Assumption	
Drug acquisition costs	Cabozantinib list price taken from BNF.[50] PAS discount provided by company.[7] Compliance estimate taken from COSMIC-311.[19]	N/a
Drug administration costs	NHS Reference Costs 2020/21[51] and PSSRU 2021[52]	N/a
Health state costs	Resource use requirements (tests, scans and visits) are based on NICE TA742.[13] Unit costs were taken from NHS Reference Costs 2020/21.[51]	
TEAE management costs	NHS Reference Costs 2020/21[51]	
End of life care costs	Georghiou and Bardsley,[53] inflated to current values using PSSRU pay and prices indices.[52]	

BSC - best supportive care; PFS - progression-free survival; OS - overall survival; TTD - time to treatment discontinuation; TTO - time trade-off; TEAE - treatment-emergent adverse event; N/a - not applicable; TA - Technology Appraisal; BNF - British National Formulary; PSSRU - Personal Social Services Research Unit; NICE - National Institute for Health and Care Excellence; RCC - renal cell carcinoma

Time-to-event parameters

Statistical adjustment of OS data to account for treatment switching

As discussed in Section 4.2.1, within both groups of the COSMIC-311 trial,[19] a change in treatment could occur following disease progression. Treatment was allowed to continue until the Investigator deemed that the patient was no longer obtaining clinical benefit or intolerable toxicity. Forty patients (45%) in the placebo arm crossed over to receive cabozantinib after progression and 11 patients (6.5%) in the cabozantinib arm continued treatment with cabozantinib after progression. The median time from progression to switching to cabozantinib for patients in the BSC arm who switched was 34 days. The company applied three treatment switching methods (inverse probability censoring weighting [IPCW],

two-stage estimation [TSE] and rank-preserving structural failure time [RPSFT] models) to adjust for placebo patients crossing over to cabozantinib treatment after progression. No adjustments were made for patients in the intervention group who continued treatment with cabozantinib after progression.

The IPCW method relies on the no unmeasured confounders assumption. The company implemented the IPCW method using in-house R routines and included time (modelled as a quadratic effect), a time-dependent progression variable (flagging patients who progressed within the next 34 days), age group and previous use of lenvatinib in the calculation of the probability of switching (see clarification response,[17] question B10).

The TSE method relies on the assumption of no unmeasured confounders at some secondary baseline. The company chose disease progression as the secondary baseline and implemented the two-stage method using in-house R routines and the *aftreg* function from the *eha* package.[54] The covariates adjusted were age group ($\leq$ 65 & > 65 years) and prior lenvatinib use (see clarification response,[17] question B11). Weibull, Gompertz, log-logistic and log-normal models were fitted, with the log-normal determined as the best-fitting model. The gamma and generalised gamma models were not used because the *eha* package does not allow for the inclusion of these two distributions (see clarification response,[17] question B11). Re-censoring was considered in the two-stage method. However, no re-censoring was applied because the estimated counterfactual time did not exceed the maximum observed survival (see clarification response,[17] question B9).

The RPSFT method relies on the common treatment effect assumption. The company used the *rpsftm* R package to implement the RPSFT method,[55] and performed sensitivity analysis to test the common treatment effect assumption. Re-censoring was applied to the RPFST method and the company reported that only 8 patients were re-censored (see clarification response,[17] question B9). In response to clarification question B12, the company provided a plot of counterfactual event times between the two treatment groups (reproduced in Figure 13) and concluded that the distributions of counterfactual event times are similar which supports the common treatment effect assumption. The company also provided the "z graph" and concluded that the estimation of psi was robust.

Figure 13: Kaplan-Meier plots of counterfactual event times (reproduced from clarification response, question B12)

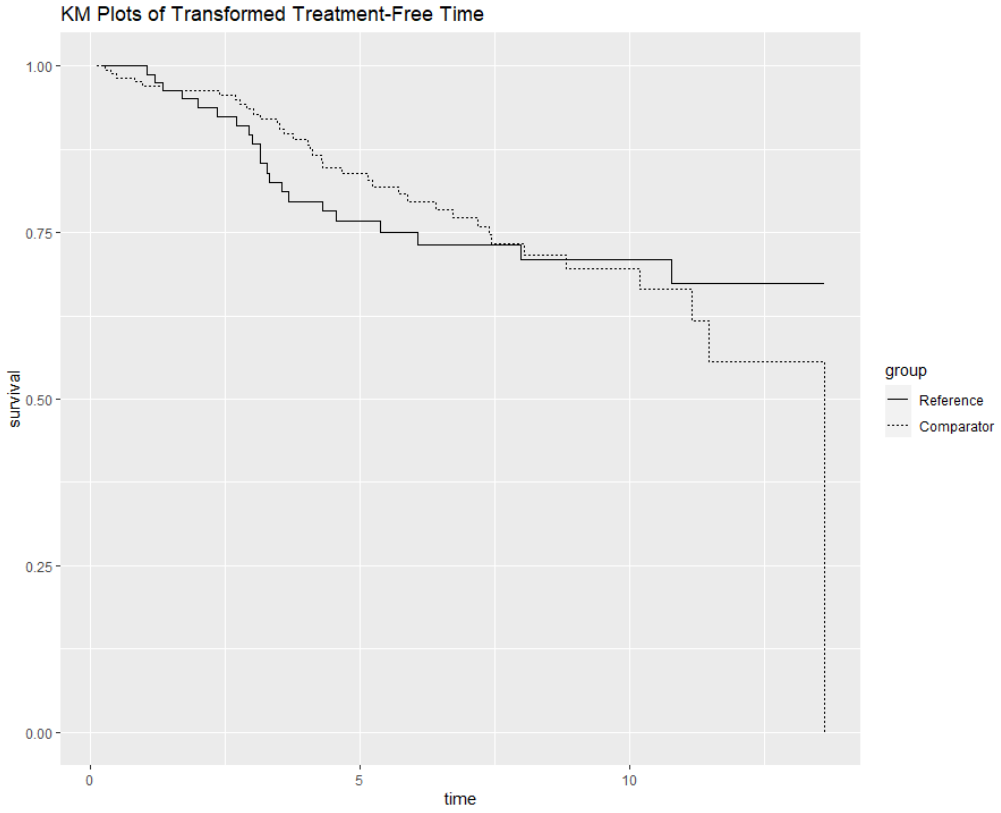
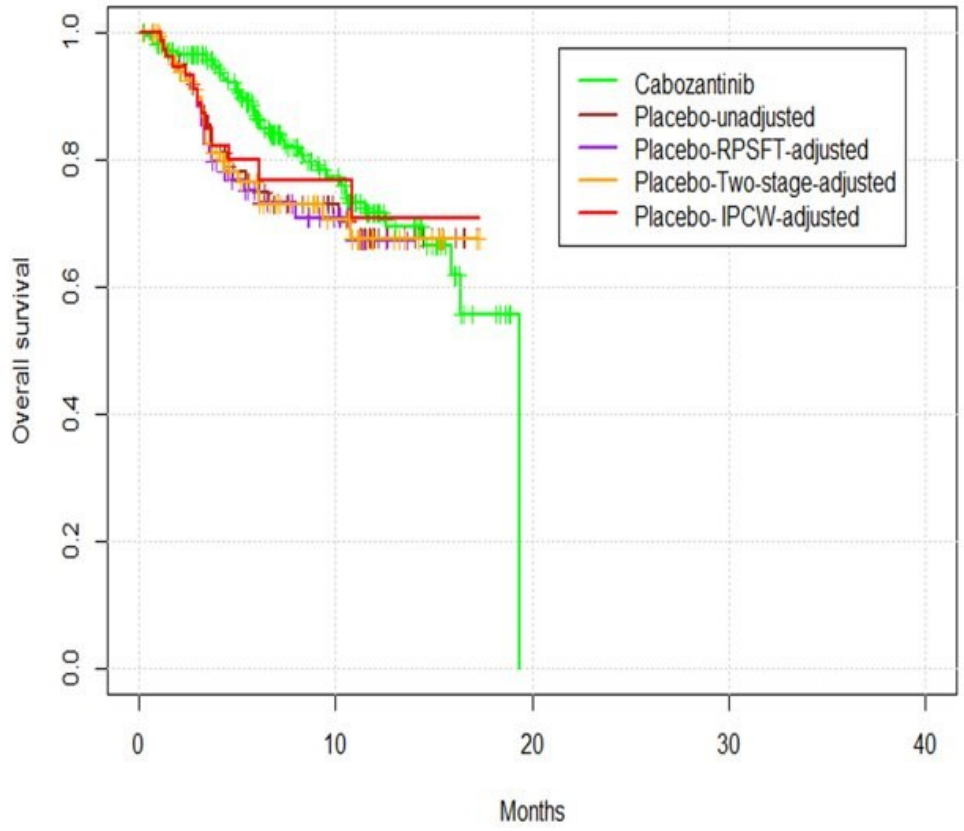


Figure 14 and Table 22 summarise the results of the three treatment switching methods implemented by the company. The two-stage-adjusted and RPSFT-adjusted placebo OS Kaplan-Meier curves were slightly lower than the unadjusted placebo OS Kaplan-Meier curve, whereas the IPCW-adjusted placebo OS Kaplan-Meier curve was slightly higher than the unadjusted placebo curve. The estimated stratified HRs from the three adjustment methods were all lower than the unadjusted HR (0.65 to 0.70 vs. 0.76). The company's sensitivity analysis assessing the common treatment effect assumption shows that the treatment effect varies from [redacted] to [redacted] when changing the assumption from the common treatment effect to no treatment effect in crossover patients (see clarification response,[17] question B14).

The RPSFT method was chosen as the base case adjustment method for the economic model based on the company's justification that it was in line with previous NICE submissions, in particular TA535.[12] In response to clarification questions B13 and B15,[17] the company provided additional justification on why the IPCW and two-stage method were not preferred over the RPSFT method. Specifically, the company noted difficulty in justifying the no unmeasured confounders assumption with these two methods because limited covariates were included in the analysis and the IPCW may not be stable as only [redacted].

Figure 14: Overall survival Kaplan-Meier curves of the three treatment switching adjustment methods (reproduced from CS, Figure 11)



RPSFT - rank preserved structural failure time; IPCW - inverse probability of censoring weights

Table 22: Overall survival results for cabozantinib vs. placebo before and after treatment switching adjustments (reproduced from CS, Table 14)

Distribution	Placebo-unadjusted (95% CI)	Placebo-RPSFT (95% CI)	Placebo-two-stage (95% CI)	Placebo-IPCW (95% CI)
Stratified HR (naïve 95% CI)	0.76 (0.45, 1.31)	0.65 (0.28, 1.53)	0.70 (0.41, 1.22)	0.68 (0.37, 1.27)
Mean survival - cabozantinib	37.58 (27.08, 50.74)	37.58 (27.08, 50.74)	37.58 (27.08, 50.74)	37.58 (27.08, 50.74)
Mean survival - placebo	30.45 (20.89, 45.71)	27.39 (18.38, 41.15)	29.25 (18.83, 43.47)	31.76 (19.5, 51.59)
Mean difference - cabozantinib vs. placebo	7.13 (-10.01, 24.27)	10.19 (-6.95, 27.33)	8.33 (-8.81, 25.47)	5.82 (-11.7, 23.34)

CI - confidence interval; IPCW - inverse probability of censoring weights; RPSFT - rank preserved structural failure time; HR - hazard ratio

Summary of parametric survival model fitting process and model selection

The company fitted a series of parametric survival models to the time-to-event data on PFS, OS (adjusted for treatment switching in placebo group) and TTD from COSMIC-311.[19] The data cut-off for all three endpoints was the 8th of February 2021 (CCO2).

The same general survival modelling approach was applied to TTD, PFS and OS (RPSFT-adjusted in the placebo group). The company fitted six standard parametric survival models to the data for each endpoint; these included the exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma distributions. Models were fitted independently to data for each treatment group without the use of a treatment indicating covariate (an HR or acceleration factor [AF]). The minutes of the company's advisory board meeting[21] indicate that the 2-parameter gamma distribution was also fitted; however, this distribution is not considered further in the CS[7] and it is not included in the company's executable model. More flexible parametric survival distributions, such as restricted cubic spline (RCS) models, were not considered.

The CS[7] states that the company's model selection process included: (i) examination of the goodness-of-fit of the models based on the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC); (ii) visual inspection of the models against the observed Kaplan-Meier survival functions and (iii) consideration of the clinical plausibility of the model predictions based on input from three UK clinicians who attended the company's 2022 advisory board meeting.[21] The company also explored how OS had been modelled in the previous NICE appraisal of sorafenib and lenvatinib for RAI-refractory DTC (TA535).[12] Hazard plots, log-cumulative hazard plots and quantile-quantile plots are not presented or discussed in the CS.

Progression-free survival

Comparisons of the observed Kaplan-Meier survival functions and parametric survival model predictions of PFS for cabozantinib and BSC are shown in Figure 15 and Figure 16, respectively. AIC and BIC statistics for the fitted models are summarised in Table 23.

For the cabozantinib group, the log-logistic distribution was the best-fitting model in terms of both AIC and BIC. For the BSC group, the generalised gamma and the log-normal distributions were the best-fitting models based on AIC and BIC, respectively. The company selected the Weibull model for both treatment groups in the base case analysis based on clinical input obtained from the experts who attended the advisory board meeting. [redacted][21] The EAG notes that for the cabozantinib group, the AIC and BIC values for the Weibull distribution are similar to those for the best-fitting (log-logistic) model; however, for the BSC group, the Weibull distribution has a noticeably worse fit than the log-normal, log-logistic and generalised gamma models.

Figure 15: Kaplan-Meier plots and parametric models, PFS, cabozantinib group (generated using the company's model)

PFS - progression-free survival; KM - Kaplan-Meier

Figure 16: Kaplan-Meier plots and parametric models, PFS, BSC group (generated using the company's model)

PFS - progression-free survival; KM - Kaplan-Meier

Table 23: AIC and BIC statistics, PFS

Model	Cabozantinib		BSC	
	AIC	BIC	AIC	BIC
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]

AIC - Akaike Information Criterion; BIC - Bayesian Information Criterion; BSC - best supportive care
** Best-fitting model indicated in bold*

Overall survival

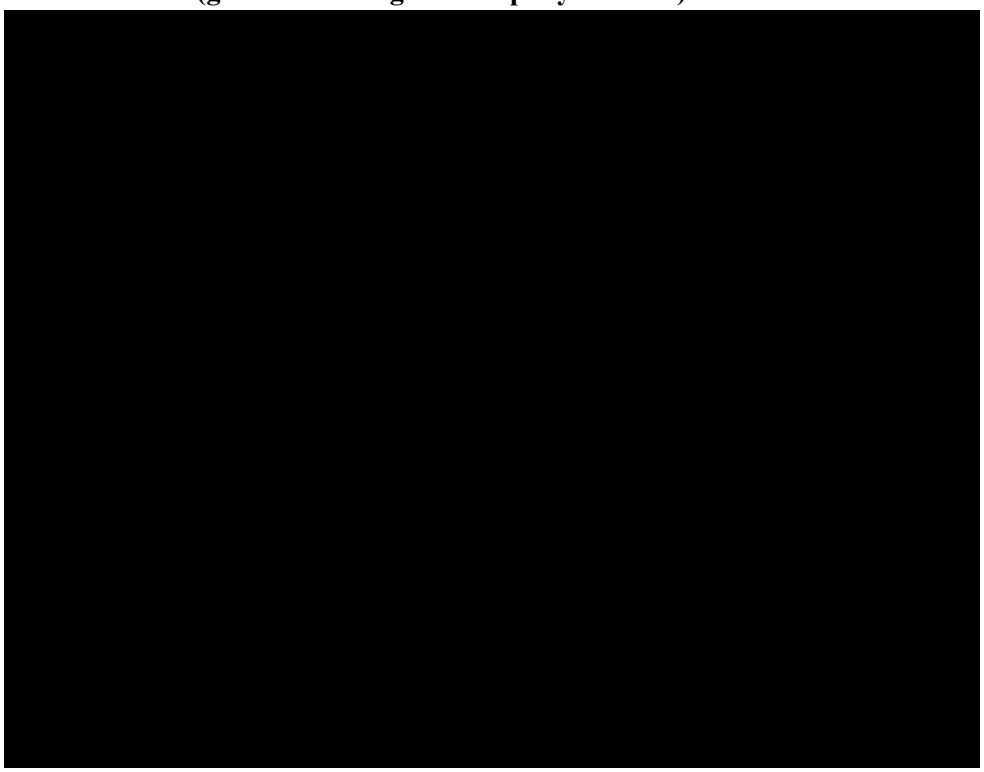
Comparisons of the observed Kaplan-Meier survival functions and parametric survival model predictions of OS for cabozantinib and RPSFT-adjusted BSC are shown in Figure 17 and Figure 18, respectively. Figure 19 presents the same information for both treatment groups on a single plot to highlight where the survival models for each treatment group cross. AIC and BIC statistics for the survival models are summarised in Table 24.

For the cabozantinib group, the Weibull distribution was the best-fitting model in terms of the AIC, although all models provided a broadly similar fit. The exponential distribution was the best-fitting model in terms of BIC; the fit was similar for the Weibull, Gompertz and log-logistic models. For BSC, the generalised gamma distribution was the best-fitting model according to both AIC and BIC; none of the other models had similar AIC values, although the BIC for the exponential and log-normal models was broadly similar. The company selected the exponential distribution for both treatment groups; this decision was influenced by clinical input[21] and was further justified through reference to the modelling approach used in TA535,[12] which applied a hybrid Kaplan-Meier function with a parametric (exponential) tail based on long-term OS data for TC patients in the Surveillance, Epidemiology, and End Results (SEER) Program. The minutes of the company's advisory board meeting provide further information regarding how clinical plausibility was used to inform model selection. [redacted]. In order to address the over-prediction of OS in the BSC group, the company's economic model applies an assumption which forces all BSC-treated patients who remain alive at 5 years to die at this timepoint. No adjustment is applied to the exponential model for OS in the cabozantinib group.

Figure 17: Kaplan-Meier plots and parametric models, OS, cabozantinib group (generated using the company's model)

OS - overall survival; KM - Kaplan-Meier
** Plot excludes general population mortality constraint*

Figure 18: Kaplan-Meier plots and parametric models, OS (RPSFT-adjusted), BSC group (generated using the company's model)



OS - overall survival; KM - Kaplan-Meier
** Plot excludes general population mortality constraint and assumption that all BSC-treated patients die by 5 years*

Figure 19: Kaplan-Meier plots and parametric models, OS (RPSFT-adjusted), both treatment groups (cabozantinib group shown as solid lines, BSC group shown as dashed lines)

OS - overall survival; KM - Kaplan-Meier
** Plot excludes general population mortality constraint and assumption that all BSC-treated patients die by 5 years*

Table 24: AIC and BIC statistics, OS

Model	Cabozantinib		BSC	
	AIC	BIC	AIC	BIC
Exponential				
Weibull				
Gompertz				
Log-normal				
Log-logistic				
Generalised gamma				

AIC - Akaike Information Criterion; BIC - Bayesian Information Criterion; BSC - best supportive care
** Best-fitting model indicated in bold*

Time to treatment discontinuation

The TTD data used in the company's model are not well-described in the CS[7] and the plots shown are subject to errors. The company's clarification response[17] (question A15) notes that the TTD analysis had missing data; hence, assumptions were required. In their response to additional clarification questions from the EAG, the company explained that TTD was analysed according to treatment end date, with the date of the subject's last known dose used for patients who were censored. An alternative analysis of TTD was also undertaken which used the censoring date for PFS instead of the last known dose; however, this alternative analysis is not included in the company's economic model and the EAG agrees with the company that the former approach is more appropriate. Despite the company's concerns

regarding missing data, the EAG is unclear why the analysis of TTD in COSMIC-311 should be considered to be any different from analyses of TTD in other oncology trials. Comparisons of the observed Kaplan-Meier survival functions and parametric survival model predictions of TTD for the cabozantinib group (before applying the PFS cap) are shown in Figure 20. AIC and BIC statistics for the fitted models are summarised in Table 25.

The exponential distribution was selected for inclusion in the company's economic model. The exponential distribution was the best-fitting model in terms of both AIC and BIC. With the exception of the log-normal distribution, all of the other models had broadly similar AIC values, and none of the other models had similar BIC values. As noted in Section 5.2.3, the model includes a cap which forces TTD to be less than or equal to PFS at all timepoints; hence, the functions shown in Figure 20 do not reflect the TTD functions applied in the model.

Figure 20: Kaplan-Meier plots and parametric models, TTD, cabozantinib group, COSMIC-311 CCO2 (Kaplan-Meier estimates digitised by the EAG from company's additional clarification response)[*†]

TTD - time to treatment discontinuation; KM - Kaplan-Meier
** The modelled TTD functions shown in the figure exclude the PFS cap*
† The Kaplan-Meier plot of TTD presented in Figure 27 of the CS does not reflect the TTD survival models as it includes the PFS cap. The plot shown in Figure 22 of the company's clarification response is incorrect as the Kaplan-Meier estimates reflect TTD linked to progression rather than TTD linked to the observed last dose

Table 25: AIC and BIC statistics, TTD

Model	Cabozantinib	
	AIC	BIC
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]
Generalised gamma	[redacted]	[redacted]

AIC - Akaike Information Criterion; BIC - Bayesian Information Criterion
** Best-fitting model indicated in bold*

Model-predicted TTD, PFS and OS

The company's base case model predictions of TTD, PFS and OS are shown in Figure 21. A summary of the predicted mean time spent in each health state is summarised in Table 26. The company's model suggests that cabozantinib extends PFS and OS compared with BSC. OS for the BSC group is assumed to drop suddenly at 5 years due to the structural assumption that all surviving patients die at this timepoint. The inclusion of a cap on TTD means that PFS and TTD are assumed to be nearly identical.

Figure 21: Model predictions of TTD, PFS and OS (generated using the company's model)

[redacted]

OS - overall survival; PFS - progression-free survival; BSC - best supportive care; TTD - time to treatment discontinuation
** Includes general population mortality constraint and assumption that all surviving BSC-treated patients die at 5 years*

Table 26: Predicted mean time in each health state (years)

Health state	Cabozantinib - mean time in state	BSC - mean time in state
Progression-free (on treatment)	[redacted]	N/a
Progression-free (off treatment)	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]
Overall survival	[redacted]	[redacted]

BSC - best supportive care; N/a - not applicable

Health-related quality of life

The COSMIC-311 trial[19] included HRQoL data collection using the EQ-5D-5L questionnaire. The questionnaire was administered at baseline, every 4 weeks for the first 25 weeks, and every 8 weeks thereafter until the later date of: (a) 8 weeks after radiographic disease progression (based on RECIST version 1.1) or (b) permanent discontinuation of study treatment. Data were not collected from placebo group patients after they switched to cabozantinib and data were only available from CCO1. The company mapped the EQ-5D-5L data from CCO1 to the 3-level (3L) version using the algorithm by Hernández Alava and Pudney.[56] The company then fitted generalised linear mixed-effect models to estimate health state utility values. Five alternative models were fitted which included a range of covariates with random intercepts and random slopes. Further details on these models can be found in the company's clarification response[17] (question B18). The full regression model ("Model 2") included covariates relating to progression status (health state), response, treatment arm, treatment status and end-of-life. The company's preferred model ("Model 1") included only progression status as a covariate and was selected for use based on statistical model fit (AIC and BIC), sample size and the requirements/capabilities of the cost-effectiveness model.[17] The CS[7] reports that the covariate for treatment arm was found not to be statistically significant; the company used this finding to justify a modelling assumption that health state utility values are independent of treatment group. Given that the preferred model did not include treatment group as a covariate, it is unclear which model this finding is based on.

As EQ-5D-5L data collection in COSMIC-311[19] stopped at disease progression, the data from the trial were not used to inform the company's model. Instead, the company used health state utility estimates from an external study – Fordham *et al*[22] – which was identified by the company's HRQoL review (see CS Appendix H,[22] Table 43). This study has been used to inform several previous NICE appraisals of treatments for TC (TA516, TA550 and TA742).[13, 57, 58] Fordham *et al.* report the methods and results of a valuation study to estimate utility values for health states associated with RAI-refractory DTC. The authors developed health state descriptions using data from a previous qualitative study[59] and through iterative review by clinical experts. The health states valued include "Best state – stable/no progression"; "response to therapy"; "progressive disease" as well as four AE-related states (diarrhoea, fatigue, hand-foot syndrome and alopecia). The health states were valued by 100 members of the general

public using the time trade-off (TTO) approach. Results are presented as observed mean utility values and as estimates derived from multivariable regression models with and without adjustment for educational qualification level and EQ-5D-3L usual activity and anxiety/depression domain scores calculated using UK norms. The company's model applies utility values of 0.87 for the progression-free state and 0.52 for the progressed disease state based on the regression model with adjustment.

The company's economic model also includes QALY losses associated with Grade 3/4 AEs, assuming a duration of 1 month. Disutility values for PPES, diarrhoea and fatigue were taken from Fordham *et al.*;[22] the disutility value for hypertension was taken from the AXIS trial of axitinib versus sorafenib for renal cell carcinoma (RCC).[23] Hypocalcaemia and proteinuria were assumed to have no impact on HRQoL. All QALY losses related to AEs were applied in the first model cycle. Overall QALY losses attributable to AEs were estimated to be -0.085 for cabozantinib and -0.004 for BSC.

The health state utility values and AE-related disutility values used in the company's model are presented in Table 27.

Table 27: Utility and disutility values used in the company's model

Health state utility values	Mean value	Source and method (population)
Progression-free	0.87	Fordham *et al.,*[22] TTO (general public)
Progressed disease	0.52	
AE disutility values		
PPES	0.34	Fordham *et al.,*[22] TTO (general public)
Diarrhoea	0.47	
Fatigue	0.08	
Hypertension	0.13	Rini *et al.*[23] EQ-5D-3L (trial patients)
Hypocalcaemia	0.00	Assumption
Proteinuria	0.00	

PPES - palmar-plantar erythrodysaesthesia syndrome; AE - adverse event; TTO - time trade-off; NR - not reported; EQ-5D-3L - Euroqol 5-Dimensions (3 Level)

The company's model includes the adjustment of utility values for increasing age. This adjustment was implemented using utility decrement multipliers for each age compared with a "source publication population age" of 67 years. The EAG notes that this age does not reflect the modelled age of the population in COSMIC-311[19] or the age of respondents or health state descriptions in the Fordham *et al.* TTO study[22] (see Section 5.3.5).

Resource use and unit costs

The model includes costs associated with: (i) drug acquisition; (ii) prescribing and dispensing of cabozantinib; (iii) health state resource use; (iv) the management of AEs, and (v) end-of-life care. The costs applied in the company's model are summarised in Table 28; individual cost components are described in further detail below.

Table 28: Summary of cost parameters used in the model

Cost component	Cabozantinib	BSC
Drug acquisition costs (per monthly cycle)*	[redacted]	£0.00
Drug administration costs (per monthly cycle)	£245.00 (month 1) £27.00 (month 2+)	£0.00
Health state costs - progression-free (per monthly cycle)	£381.96	£354.88
Health state costs - progressed disease (per monthly cycle)	£268.86	
AE management costs (once-only)	£191.58	£28.66
End-of-life care (once-only)	£8,705.50	

** Includes adjustment for packs per month and compliance*
AE - adverse event; BSC - best supportive care; PAS - Patient Access Scheme

Drug acquisition and administration costs

Drug acquisition costs for cabozantinib are summarised in Table 29. The list price for cabozantinib is £5,143.00 per pack of 30 x 60mg tablets. The company has an agreed PAS which takes the form of a simple price discount of [redacted]; including this discount results in a cost per pack of [redacted]. Within the model, acquisition costs are calculated as a function of the 60mg daily dosing schedule for cabozantinib,[20] the TTD in COSMIC-311 (capped by PFS),[19] compliance and the PAS price for cabozantinib. The model does not include any costs associated with drug wastage. The model applies a drug acquisition cost of [redacted] during each month in which the patient remains on progression-free and treatment.

Table 29: Cabozantinib acquisition costs

Parameter	Value	Source
Cabozantinib list price (30 tablet pack)	£5,143.00 (list price) [redacted]	BNF[50]
Packs per month	1.01	-
PAS discount	[redacted]	CS[7]
Compliance*	[redacted]	COSMIC-311[19]
Cost per month	[redacted]	-

PAS - Patient Access Scheme; BNF - British National Formulary; CS - company's submission
** Calculated as the proportion of days in the trial in which treatment was received*

Administration costs for cabozantinib are assumed to include the cost of delivering chemotherapy from NHS Reference Costs 20/21 (SB11Z - deliver exclusively oral chemotherapy)[51] in the first model cycle and an additional 30 minutes of pharmacists' time from the Personal Social Services Research Unit (PSSRU)[52] during each subsequent cycle in which the patient remains progression-free and on treatment.

The economic model does not include any concomitant drug therapy costs related to treatment with BSC.

Health state resource use

Resource costs related to disease management include the costs associated with medical visits (consultants and nurses), tests and imaging (different types of blood tests, CT scans and ECGs). These costs are assumed to be independent of treatment group, with the exception of ECGs in the progression-free state which are included for cabozantinib but not for BSC. Resource use estimates were based on NICE TA742,[13] with unit costs valued using NHS Reference Costs 2020/21.[51] Table 30 presents the per-cycle costs for the progression-free and progressed disease health states in the company's model.

Table 30: Health state resource use and costs (monthly)

Resource use item	Resource use per month			Unit cost	Costs per monthly cycle			Unit cost source
	PF - cabo	PF - BSC	PD - both groups		PF - cabo	PF - BSC	PD - both groups	
Blood test routine U&Es	1.00	1.00	0.50	£1.85	£1.85	£1.85	£0.93	DAPS04, clinical biology
Haematology/ coagulation test	1.00	1.00	0.50	£3.63	£3.63	£3.63	£1.82	DAPS05, clinical haematology
Blood test calcium and magnesium	1.00	1.00	0.50	£1.85	£1.85	£1.85	£0.93	DAPS04, clinical biochemistry
LFT	1.00	1.00	0.50	£1.85	£1.85	£1.85	£0.93	DAPS04, clinical biochemistry
TFT	1.00	1.00	0.50	£1.85	£1.85	£1.85	£0.93	DAPS04, clinical biochemistry
Consultant-led outpatient visits	1.00	1.00	0.50	£224.55	£224.55	£224.55	£112.28	Consultant-led, non-admitted face-to-face attendance, follow up (Medical Oncology - 370/WF01A)
Nurse-led outpatient visits	0.33	0.33	0.50	£190.59	£63.53	£63.53	£95.30	Non-consultant-led, non-admitted face-to-face attendance, follow up (medical oncology - 370/WF01A)
CT scan	0.33	0.33	0.33	£167.31	£55.77	£55.77	£55.77	CT scan of more than 3 areas (RD27Z)
ECG	0.17	0.00	0.00	£162.46	£27.08	£0.00	£0.00	Outpatient procedures. medical procedures (EY51Z)
Total cost	**-**	**-**	**-**	**-**	**£381.96**	**£354.88**	**£268.86**	-

Cabo - cabozantinib; BSC - best supportive care; PF - progression-free; PD - progressed disease; U&E - urea and electrolytes; LFT - liver function test; TFT - thyroid function test; CT - computerised tomography; ECG - electrocardiogram

AE management costs

Costs related to the management of AEs were based on the frequency of individual Grade 3/4 TEAEs with an incidence ≥5% observed in either the cabozantinib arm or placebo arm of the ITT population of the COSMIC-311 trial at CCO1.[19] Unit costs were taken from NHS Reference Costs 2020/21[51] and assumptions. AE frequencies, unit costs and total costs used in the model are summarised in Table 31. AE management costs per patient for cabozantinib and BSC are estimated to be £191.58 and £28.66, respectively. These costs are applied once-only during the first model cycle.

Table 31: Adverse event costs

AE	Frequency		Unit cost	Event cost		Unit cost source
	Cabo	BSC		Cabo	BSC	
Hand-foot syndrome	10.00%	0.00%	£490.67	£49.07	£0.00	JD07K. CC Score 0-1. NES (Total HRGs)
Proteinuria	1.00%	0.00%	£224.55	£2.25	£0.00	Medical oncology - 370/WF01A
Hypertension	9.00%	3.00%	£537.86	£48.41	£16.14	EB04Z NES (total HRGs)
Diarrhoea	7.00%	0.00%	£635.99	£44.52	£0.00	FD10M. CC score 0-2. NES (Total HRGs)
Fatigue	8.00%	0.00%	£44.00	£3.52	£0.00	PSSRU. Community-based - nurse unit cost (including qualifications)
Hypocalcaemia	7.00%	2.00%	£625.96	£43.82	£12.52	SA09L CC Score 0-1. NES (Total HRGs)
Total cost	-	-	-	**£191.58**	**£28.66**	-

AE - adverse event; cabo - cabozantinib; BSC - best supportive care; NES - non-elective short stay; CC - complications and comorbidities; CL - consultant-led; PSSRU - Personal Social Services Research Unit; HRG – Healthcare Resource Group

End-of-life care costs

The cost of end-of-life care was estimated to be £8,705.50 per patient, which is applied as a once-only cost to patients at the point of death. This cost was based on Georghiou and Bardsley,[53] and was assumed to include the costs of care with General Practitioner (GP) and district nurse visits, social care, inpatient admissions, outpatient attendances and Accident and Emergency (A&E) visits at the end of life for patients with DTC. The reported estimates were uplifted to 2021 values using the NHS Cost Inflation Index (NHSCII) and Hospital and Community Health Services (HCHS) indices.[52]

5.2.5 Model evaluation methods

The CS[7] presents base case cost-effectiveness results for cabozantinib versus BSC using the using both the deterministic and probabilistic versions of the model. The probabilistic incremental cost-effectiveness ratio (ICER) is based on 10,000 Monte Carlo simulations. The results of the probabilistic sensitivity analysis (PSA) are presented using a cost-effectiveness plane and cost-effectiveness acceptability curves (CEACs). The distributions used in the company's PSA are summarised in Table 32.

The CS[7] presents the results of the deterministic sensitivity analyses (DSAs) graphically using a tornado plot and in tabular form. The CS also reports on nine scenario analyses exploring alternative discount rates, the exclusion of age-adjustment of utility values and alternative parametric survival distributions for PFS and OS. These scenario analyses are presented using both the deterministic and probabilistic versions of the model.

The CS[7] reports an estimated value of the decision modifier calculated using the York QALY shortfall calculator,[60] based on the average age and sex of the modelled cohort and the mean discounted QALYs predicted for the BSC group. The results presented in the CS do not include QALY weighting; instead, the CS (page 130) reports a cost-effectiveness threshold range for decision-making which is adjusted according to the severity weighting.

Table 32: Distributions used in company's PSA

Parameter / group	Distribution	EAG comments
Start age	Gamma	-
Probability male	Beta	Applies arbitrary SE of 10% rather than SE from COSMIC-311
PFS	Multivariate normal	RPSFTM-adjusted OS estimates for BSC are treated as observed data, ignoring additional uncertainty associated with the switching analysis. The company's probabilistic sampling sub-routine returns errors for a proportion of samples for the generalised gamma and log-normal distributions. This appears to be caused by invalid samples of survival model coefficients.
OS and TTD	Normal	
General population mortality	Fixed	-
Health state utility values	Fixed	These parameters are uncertain and should have been included in the PSA.
AE total QALY loss	Beta	Aggregate QALY loss sampled - uncertainty around underlying parameters (AE frequencies and disutilities) is not sampled. Applies arbitrary SE of 10%.
AE total cost	Beta	Applies arbitrary SE of 10%.
Drug acquisition costs	Fixed	-
Drug administration costs	Fixed	These parameters are uncertain and should have been included in the PSA.
Health state costs	Gamma	Aggregate health state costs sampled - uncertainty around underlying parameters (resource use frequency and unit costs) is not sampled. Applies arbitrary SE of 10%.
TEAE management costs	Gamma	Applies arbitrary SE of 10%.
End of life care costs	Gamma	Applies arbitrary SE of 10%.

EAG - External Assessment Group; SE - standard error; PFS - progression-free survival; OS - overall survival; TTD - time to treatment discontinuation; AE - adverse event; QALY - quality-adjusted life year; PSA - progression-free survival

5.2.6 Company's original model results

Company's central estimates of cost-effectiveness

Table 33 presents the central estimates of cost-effectiveness generated using the company's original model. All results include the agreed PAS for cabozantinib and exclude QALY weighting (unless otherwise stated). The probabilistic version of the model suggests that cabozantinib is expected to generate an additional [redacted] discounted QALYs at an additional cost of [redacted]; the corresponding ICER is £27,169 per QALY gained. The deterministic version of the model results in a slightly higher ICER of £28,148 per QALY gained. The base case analysis suggests a decision modifier of 1.2 (age = 65 years; 53% female; 1.19 discounted QALYs for the comparator group). When QALY weighting is included, the probabilistic version of the company's model suggests an ICER of £22,641 per QALY gained.

Table 33: Company's original base case model results, cabozantinib versus BSC (generated by the EAG, excluding QALY weighting)

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER	DM
Probabilistic model†								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£27,169**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Deterministic model								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£28,148**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-

LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; BSC - best supportive care; DM - decision modifier

** Undiscounted*

† Based on a re-run of the company's PSA sub-routine by the EAG

Company's PSA results

Figure 22 presents the results of the company's PSA in the form of CEACs for cabozantinib and BSC. Excluding QALY weighting, the probability that cabozantinib generates more net benefit than BSC at willingness-to-pay (WTP) thresholds of £20,000 and £30,000 per QALY gained is approximately [redacted] and [redacted], respectively. When QALY weighting is included, the corresponding probabilities are [redacted] and [redacted], respectively.

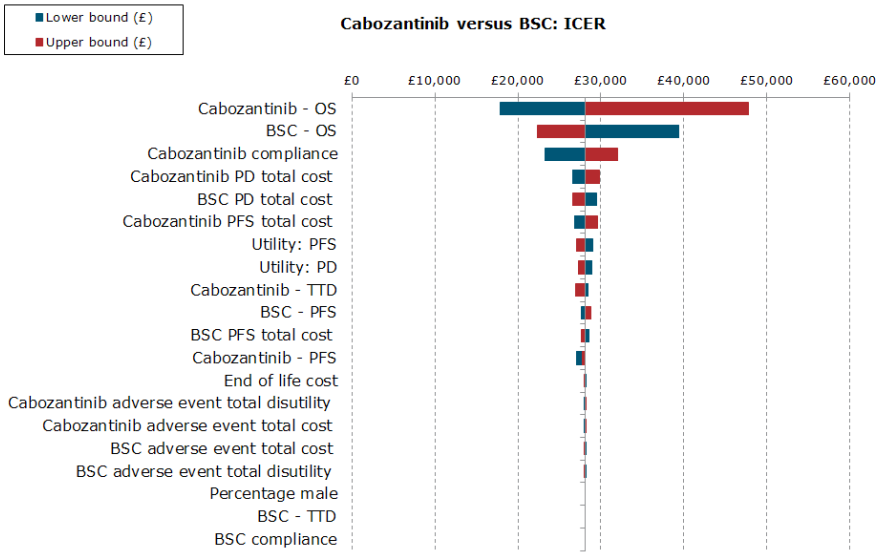
Figure 22: CEACs, cabozantinib versus BSC (excluding QALY weighting)

BSC - best supportive care

Company's DSA results

Figure 23 presents the results of the company's DSAs in the form of a tornado plot. The plot indicates that the ICER is particularly sensitive to the OS rate parameter in both the cabozantinib and BSC groups (which drives total QALYs) and the probability of compliance with cabozantinib (which influences net drug acquisition costs). Across the range of scenarios presented, the ICER ranges from £17,920 to £47,776 per QALY gained (excluding QALY weighting).

Figure 23: Company's DSA results, cabozantinib versus BSC (generated using the company's model, excluding QALY weighting)



BSC - best supportive care; ICER - incremental cost-effectiveness ratio; OS - overall survival; PD - progressed disease; TTD - time to treatment discontinuation; PFS - progression-free survival

Company's scenario analyses

Figure 23 presents the results of the company's scenario analyses. As shown in the table, the ICER is not sensitive to the discount rate, the inclusion/exclusion of age-adjustment of utility values or the selected PFS model. The ICER is lower for the log-normal OS model, although no other models have been explored in the scenario analyses presented in the CS[7] (the EAG presumes this is because no other fitted models were considered plausible). Amongst the scenarios considered by the company, the highest ICER reported is £30,567 per QALY gained (PFS = exponential model). The CS also reports the results of these scenario analyses generated using the probabilistic version of the model. The results presented in the CS suggest noticeable differences between the deterministic and probabilistic scenario analyses. However, the EAG re-ran all of the company's probabilistic scenario analyses and found the results to be generally similar to their deterministic counterparts. The reasons for the apparent discrepancies in the company's results are unclear.

Table 34: Company's scenario analyses (generated by the EAG, excluding QALY weighting)

Scenario no.	Scenario	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER	DM
Deterministic model						
-	Company's base case	[redacted]	[redacted]	[redacted]	£28,148	1.2
S1	Discount rate - 0%	[redacted]	[redacted]	[redacted]	£26,165	1.2
S2	Discount rate - 5%	[redacted]	[redacted]	[redacted]	£28,976	1.2
S3	Utility not age-adjusted	[redacted]	[redacted]	[redacted]	£27,937	1.2
S4	PFS - exponential	[redacted]	[redacted]	[redacted]	£30,567	1.2
S5	PFS - generalised gamma	[redacted]	[redacted]	[redacted]	£29,937	1.2
S6	PFS - Gompertz	[redacted]	[redacted]	[redacted]	£27,848	1.2
S7	PFS - log-logistic	[redacted]	[redacted]	[redacted]	£27,740	1.2
S8	PFS - log-normal	[redacted]	[redacted]	[redacted]	£27,718	1.2
S9	OS - log-normal	[redacted]	[redacted]	[redacted]	£19,617	1.2
Probabilistic model						
-	Company's base case	[redacted]	[redacted]	[redacted]	£27,169	1.2
S1	Discount rate = 0%	[redacted]	[redacted]	[redacted]	£25,065	1.2
S2	Discount rate = 5%	[redacted]	[redacted]	[redacted]	£27,901	1.2
S3	Utility not age-adjusted	[redacted]	[redacted]	[redacted]	£26,821	1.2
S4	PFS = exponential	[redacted]	[redacted]	[redacted]	£28,267	1.2
S5	PFS = generalised gamma*	[redacted]	[redacted]	[redacted]	£25,386	1.2
S6	PFS = Gompertz	[redacted]	[redacted]	[redacted]	£17,592	1.2
S7	PFS = log-logistic	[redacted]	[redacted]	[redacted]	£26,247	1.2
S8	PFS = log-normal*	[redacted]	[redacted]	[redacted]	£25,161	1.2
S9	OS = log-normal*	[redacted]	[redacted]	[redacted]	£16,961	1.2

S - scenario; LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; DM - decision modifier; PFS - progression-free survival; OS - overall survival

** For these scenarios, samples had to be deleted from the final PSA results as they returned #NUM! errors (5 samples, removed for Scenario S5, 50 samples removed for Scenario S8, 682 samples removed for Scenario S9).*

5.3 Critical appraisal

This section details the EAG's critical appraisal of the company's original economic model. Following the clarification round, the company submitted a revised model which includes a number of amendments to the model assumptions and parameters. This amended model is described and critiqued separately in Section 5.4.

5.3.1 Critical appraisal methods

The EAG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic analysis and the underlying health economic model upon which this is based. These included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists.[61, 62]
- Scrutiny and discussion of the company's model by the EAG.
- Double-programming of the deterministic version of the company's model to fully assess the logic of the model structure, to draw out any unwritten assumptions and to identify any apparent errors in model implementation.
- Examination of the correspondence between the description of the model reported in the CS[7] and the company's executable model.
- Replication of the base case results, PSA, DSAs and scenario analyses reported in the CS using the company's executable model.
- Where possible, checking of key parameter values used in the company's model against their original data sources.
- The use of expert clinical input to judge the credibility of the company's economic analyses and the assumptions underpinning the model.

5.3.2 Model verification by the EAG

The EAG rebuilt the deterministic version of the company's base case model in order to verify its implementation. As shown in Table 35, the EAG's results are very similar to those generated using the company's original model. During the process of rebuilding the model, the EAG identified several minor programming errors; these are described in detail in Section 5.3.5 (critical appraisal point [1]). The correction of these errors forms part of the EAG's exploratory analyses (see Section 5.5).

Table 35: Comparison of results from the company's original base case model and the EAG's double-programmed model (excluding the correction of errors identified by the EAG)

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. Costs	ICER
Company's deterministic model							
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£28,148**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-
EAG's double-programmed model							
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£28,150**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-

LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; BSC - best supportive care
** Undiscounted*

5.3.3 Correspondence of the model inputs and the original sources of parameter values

Where possible, the EAG checked the company's model input values against their original sources. The EAG was able to identify the utility/disutility values, the AE frequencies, the resource use estimates and the unit costs used in the company's model. The majority of the other model parameters, including the survival model parameters, were generated from analyses of IPD from the COSMIC-311 trial.[19] These data were not made available to the EAG; hence, the EAG is unable to verify that the analyses have been undertaken appropriately.

5.3.4 Adherence to NICE Reference Case

Table 36 summarises the extent to which the company's economic model adheres to the NICE Reference Case.[63] Overall, the EAG believes that the company's model is generally in line with the Reference Case. The most pertinent deviation relates to the use of utility values obtained from a general population TTO study in preference to the EQ-5D-5L data collected in COSMIC-311.[19] This issue is discussed further in Section 5.3.5, critical appraisal point [5].

Table 36: Adherence to the NICE Reference Case

Element of HTA	Reference Case	EAG comments
Defining the decision problem	The scope developed by NICE	The company's economic analysis is in line with the final NICE scope.[18] The model compares cabozantinib versus BSC. However, the EAG's clinical advisors commented that some patients continue to receive lenvatinib following disease progression. No clinical or economic comparison has been presented between cabozantinib and continued post-progression lenvatinib.
Comparator(s)	As listed in the scope developed by NICE	
Perspective on outcomes	All health effects, whether for patients or, when relevant, carers	The model includes health outcomes accrued by patients. Health impacts on caregivers are not included.
Perspective on costs	NHS and PSS	Costs reflect those borne by the NHS and PSS
Types of economic evaluation	Cost-utility analysis with fully incremental analysis	The model is evaluated using a cost-utility approach.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The model includes a 35-year (lifetime) horizon. At the end of the time horizon, virtually all (>99.99%) patients in both treatment groups have died.
Synthesis of evidence on health effects	Based on systematic review	Health outcomes are modelled using data collected in the COSMIC-311 trial.[19] This is the pivotal Phase 3 placebo-controlled trial of cabozantinib for RAI-refractory DTC. This study was identified in the company's SLR.[7, 25]
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults	Health state utility values are taken from a general population TTO valuation study of RAI-refractory DTC health states reported by Fordham *et al.*[22] Whilst EQ-5D-5L data were collected in COSMIC-311,[19] these have not been used in the company's model. Disutilities associated with AEs are taken from the Fordham *et al.* TTO study, except for the disutility for hypertension which was based on EQ-5D-3L values reported in the AXIS trial[23] (lenvatinib with everolimus for advanced RCC). These disutility values have been used in previous NICE appraisals (TA498[64] and TA535[12]).
Source of data for measurement of HRQoL	Reported directly by patients or carers, or both	
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit, except in specific circumstances	The company has generated estimates of QALY shortfall which suggest a decision modifier of 1.2.
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	Unit costs are taken from NHS Reference Costs,[51] the PSSRU,[52] the BNF[50] and Georghiou and Bardsley.[53] Costs are valued at 2020/21 prices.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Health outcomes and costs are discounted at a rate of 3.5% per year.

HTA - health technology assessment; EAG - External Assessment Group; NICE - National Institute for Health and Care Excellence; NHS - National Health Service; PSS - Personal Social Services; SLR - systematic literature review; EQ-5D-5L - Euroqol 5-Dimensions 5-Level; RAI - radioactive iodine; DTC - differentiated thyroid cancer; TTO - time trade-off; TA - Technology Appraisal; HRQoL - health-related quality of life; QALY - quality-adjusted life year; PSSRU - Personal Social Services Research Unit; BNF - British National Formulary; RCC - renal cell carcinoma

5.3.5 Main issues identified from the EAG's critical appraisal

Box 1 summarises the main issues identified within the EAG's critical appraisal of the company's original economic analyses. These issues are discussed in further detail in the subsequent sections.

Box 1: Main issues identified from the critical appraisal

(1) Model errors
(2) Absence of an economic comparison against continued lenvatinib
(3) Concerns regarding company's adjustment for treatment switching
(4) Concerns regarding company's survival analysis
(5) Concerns regarding health state utility values
(6) Concerns regarding resource use and cost assumptions
(7) Weak characterisation of uncertainty

(1) Model errors

The EAG's double-programming exercise described in Section 5.3.2 revealed six minor errors in the company's original model. These are summarised below:

(i) *Incorrect half-cycle correction approach.* The half-cycle correction is applied by reading off cumulative survival probabilities of TTD, PFS and OS at cycle 0, 0.5, 1.5 etc. This approach is incorrect as it will overestimate the contribution of the first model cycle to the overall estimates of health outcomes and costs. The half-cycle correction should have been applied by taking the average of the modelled cumulative survival probabilities between consecutive cycles for each endpoint. TTD should not be half-cycle corrected.

(ii) *Use of arbitrary values to represent general population mortality risk.* The model includes a table of values which the EAG presumes was intended to reflect estimates of general population mortality risk conditional on the probability of an individual surviving up to each age (column "*qx*" in Office for National Statistics [ONS] life tables). However, the values used in the model are arbitrary numbers which increase by 0.01 in each year. ONS life tables for England should have been used instead.

(iii) *Inappropriate assumptions underpinning per cycle general population mortality risks*. The company's general population mortality risk calculations assume that: (a) men and women have different risks of death each year, and that (b) the proportion of men and women alive remains constant in every cycle. Both assumptions cannot simultaneously be true. The EAG believes that it would be more appropriate to estimate general population mortality risk using survival models for men and women weighted by their respective proportions at baseline in COSMIC-311.[19]

(iv) *Inappropriate application of mortality constraint*. The general population mortality constraint ensures that the modelled OS function for people with DTC cannot be higher than that for the

general population. The EAG believes that this constraint should have instead been applied to the per-cycle risk of death in each cycle, rather than the overall survival function.

(v) *Incorrect implementation of age-adjustment of utility values and absence of a general population utility cap*. The company's application of age-adjusted utility values results in values which are higher than EQ-5D-3L estimates for the general population. This implies that it is better to have the disease than to not have the disease. In addition, the company's utility age-adjustment calculations result in even higher values for the first year of the model due to the use of an arbitrary "source publication population age" of age 67 years. The EAG believes that it would be more appropriate to apply a cap to prevent the utility values in the model from exceeding the general population EQ-5D and for the age-adjustment calculations to reflect the decline in EQ-5D-3L for a population which is consistent with the start age in the model (65 years).

(vi) *Unnecessary use of rounding*. AE frequencies have been unnecessarily rounded to integer values.

The company's clarification response[17] (question B17) confirms that these issues are errors. The company's revised model provided as part of their clarification response includes amendments to address all of these issues, except for issue (iii). The EAG has concerns regarding the appropriateness of some of the other more substantial amendments applied in the company's revised model (see Section 5.4). The EAG's exploratory analyses include the correction of these errors (see Section 5.5).

(2) Absence of an economic comparison against continued lenvatinib

BSC is the sole comparator included in the CS[7] - this is in line with the final NICE scope.[18] The clinical experts consulted by the company and the EAG stated that lenvatinib is the preferred first-line TKI used in usual practice and that some clinicians continue to offer lenvatinib to patients who are still obtaining clinical benefit after they have progressed.[21] The EAG's clinical advisors both stated that they do not continue first-line treatment beyond progression, although one advisor stated that they might offer radiotherapy and continue lenvatinib if progression is restricted to a single site; the other advisor commented that they would not continue treatment in this clinical scenario. One of the EAG's advisors stated that they expect a roughly even split in terms of the number of clinicians who do and do not continue treatment beyond progression. In principle, if clinicians would switch from lenvatinib to cabozantinib at the point of progression, this may suggest that continued post-progression lenvatinib use could be considered as a potential comparator for second-line cabozantinib. This comparison has not been presented in the CS.

The company's clarification response[17] (question A13) states that BSC is the only relevant comparator defined in the final NICE scope.[18] The company's response also explains that continued lenvatinib or sorafenib given after progression have not been included as comparators in the CS[7] as there is no clinical evidence to inform these comparisons. The company's response also highlights that the SELECT trial[42]

(the pivotal placebo-controlled RCT of lenvatinib for DTC) did not allow lenvatinib to be given after radiological disease progression. The company's response further highlights that cabozantinib is likely to offset costs of continued lenvatinib, which might suggest that the ICER estimated by the company's model is an overestimate. Overall, the EAG agrees that it is unlikely that sufficient evidence exists to compare cabozantinib against continued post-progression lenvatinib, but notes that the restricting the comparator to BSC alone does not fully reflect clinical practice.

(3) Concerns regarding company's adjustment for treatment switching

The EAG notes that there are some limitations associated with the company's treatment switching analysis. The EAG agrees with the limitations listed by the company for the IPCW and two-stage method - only a limited set of covariates were adjusted for in these two methods and it is unlikely that the no unmeasured confounders assumption will hold. However, the EAG also questions the validity of the RPSFT approach. The EAG disagrees with the company's conclusion that the g-estimation produced a robust outcome because the counterfactual survival times between the two arms shown in Figure 13 are not very similar and there is a marked difference between the overall treatment effect estimated using the RPSFT method (AF=1.42) and the treatment effect estimated for the BSC group (AF=1.09 using the company's choice of best fitting model). This indicates the potential violation of the common treatment effect assumption. The EAG's clinical advisors stated that it is difficult to judge the common treatment effect assumption but it could be considered to be appropriate. The EAG also notes that the RPSFT-adjusted data are not very different from the unadjusted data (see Figure 14) and the adjusted and unadjusted HRs are similar (see Table 22).

The company only presented the RPSFT results with re-censoring. The RPSFT method without re-censoring was not performed. Re-censoring only impacted on the survival time for eight patients; as such, excluding re-censoring is unlikely to have a large impact on the estimation of the treatment effect.

The EAG also advises caution regarding the interpretation of the estimated stratified HRs because the naïve 95% CI was calculated instead of using bootstrapping for the IPCW and two-stage methods, and for the RPSFT method, the SE was artificially inflated to preserve the ITT *p*-values.

The company did not adjust the OS for patients who continued treatment with cabozantinib after progression in the cabozantinib arm; the CS[7] does not provide any justification for this. The EAG believes that if treatment would be discontinued at progression, as is implied by the cap applied to TTD in the company's economic model, OS for patients who continued treatment with cabozantinib should also be adjusted for. However, because this only applies to 11 patients (6.5%) in the cabozantinib group, the impact of not adjusting the OS for these patients is likely to be small.

(4) Concerns regarding company's survival analysis

The EAG has several concerns regarding the parametric survival modelling presented in the CS[7] and the assumptions applied in the company's economic model. These concerns are discussed below based on the general considerations around model fitting and selection set out in NICE Decision Support Unit (DSU) Technical Support Documents (TSDs) 14 and 21.[65, 66]

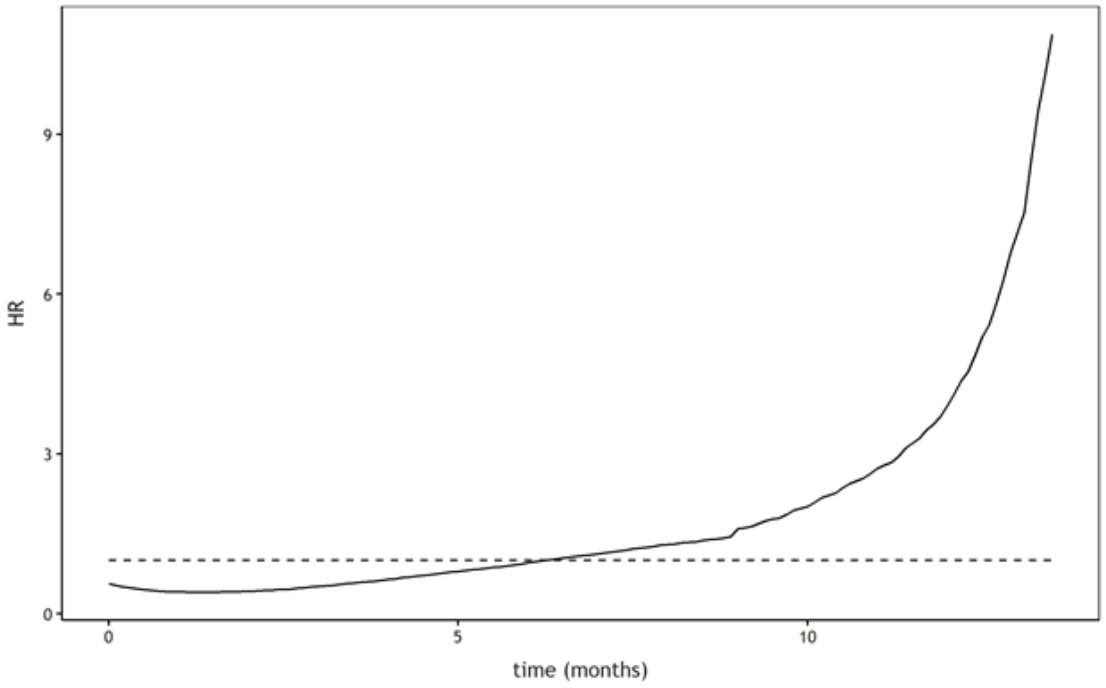
(a) Use of independent versus jointly fitted models

The CS[7] does not present any exploration around the appropriateness of applying an assumption of proportional hazards (PH) or using constant AFs to characterise relative treatment effects. As part of their clarification response, the company provided log-cumulative hazard plots, plots of Schoenfeld residuals and quantile-quantile plots (see clarification response,[17] questions B3 and B4). The log-cumulative hazard plots show that the curves cross and the plots of Schoenfeld residuals suggest non-zero slopes of the scaled residuals against time. The company also notes that the quantile-quantile plots do not indicate 45-degree lines. Taken together, the company concluded that assumptions of constant treatment effects, either in terms of HRs or AFs, are not appropriate given the observed data. The company's clarification response to question B3 highlights that the company's model is informed by independently fitted parametric models.

However, whilst the company fitted models independently to the time-to-event data for each treatment group, the model uses exponential distributions for OS in which the relative hazard is constant over time (until the BSC 5-year death assumption takes effect). This approach is equivalent to assuming PH for OS. The EAG also notes that the quantile-quantile plots used by the company to assess the constant AF assumption are incorrect because they only assess whether the exponential model is appropriate. The quantile of each treatment group should have been plotted to assess the constant AF assumption.

As part of their clarification response[17] (question B7), the company provided a plot of the time-varying HR for OS for cabozantinib versus placebo in COSMIC-311[19] (see Figure 24). The time-varying HR is increasing over time (indicating a lower treatment effect) and crosses unity (HR=1.0) after around 6 months, indicating a higher hazard in the cabozantinib group compared with the BSC group after this timepoint. The company's clarification response argues that the follow-up in COSMIC-311 is short, that some of the patterns shown in the data are not considered clinically plausible and that this strengthens the argument for using external data sources. However, the EAG notes that the company's model only uses survival data from COSMIC-311 and the use of a constant HR is inconsistent with what has been observed in the trial.

Figure 24: Empirical time-varying HR for OS (reproduced from clarification response, question B7, Figure 21)



(b) Range of models assessed

The company fitted six standard parametric models to the available data on PFS and OS (shown previously in Figure 15, Figure 16, Figure 17 and Figure 18). On the basis of visual inspection alone, the EAG notes that none of the standard parametric models provide a good representation of the observed OS data for the BSC group. The observed data for BSC indicate that the hazard of death is decreasing in the tail and only the Gompertz and generalised gamma models reflect this characteristic; however, both of these models suggest implausible extrapolations whereby more than 40% of BSC-treated patients are predicted to survive beyond 15 years (green and grey dashed lines in Figure 19). The use of more flexible parametric models may have been better able to reflect the observed data, although it is likely that this approach would also result in implausibly long tails in the OS function for the BSC group.

During the clarification process, the EAG asked the company to explore whether more flexible parametric distributions such as RCS models could provide more clinically plausible predictions of OS for cabozantinib and BSC (see clarification response,[17] question B1). However, the company did not explore these models. The company's response states that these models *"would not offer an improvement to the current limitations in the model, and are therefore not deemed appropriate."* The company's response highlights that the company would also anticipate that the OS functions for cabozantinib and BSC would follow the observed data more closely, thereby producing unrealistic predictions due to the curves crossing. The EAG agrees that this is likely, but believes that it would

have been useful to explore this set of flexible models, even if additional assumptions are required to extrapolate beyond the observed data.

(c) Statistical and visual goodness-of-fit

PFS (Figure 15, Figure 16 and Table 23). The company selected a Weibull model to estimate PFS in both groups. For the cabozantinib group, the best-fitting model is the log-logistic distribution; the AIC and BIC values for the Weibull model are similar to those for the log-normal model. For the BSC group, the best-fitting models are the generalised gamma and log-normal distributions; the AIC and BIC values for the Weibull model are markedly worse. In terms of visual inspection, the Weibull models appear to provide a reasonable representation of the observed data for both groups.

OS (Figure 17, Figure 18 and Table 24). The company used an exponential distribution to model OS in both groups. For the cabozantinib group, the exponential distribution is the best-fitting model according to the BIC and is not markedly different to the best-fitting (Weibull) model according to the AIC. For the BSC group, the generalised gamma is the best-fitting model according to both AIC and BIC; the BIC for the exponential model is similar to that for the generalised gamma model but the AIC is markedly worse. In terms of visual inspection, the selected exponential model appears to over-predict the tail of the OS function for the cabozantinib group and suggests a sharper decline in survival for the BSC group compared with the Kaplan-Meier estimate.

(d) Consideration of nature of hazards

The CS[7] does not present plots of the empirical and/or modelled hazard functions for any of the time-to-event endpoints. These plots can be useful for assessing whether the hazard functions for the selected models are consistent with the underlying empirical hazards in the observed data.

As part of the company's clarification response[17] (question B2), the company provided plots of unsmoothed, smoothed and modelled hazards for PFS and OS for the cabozantinib and RPSFT-adjusted placebo groups of COSMIC-311. The smoothed and modelled hazards for PFS for cabozantinib and placebo are shown in Figure 25 and Figure 26, respectively. The smoothed and modelled hazards for OS for cabozantinib and placebo are shown in Figure 27 and Figure 28, respectively. Unsmoothed hazard plots can be found in the company's clarification response[17] (question B2, Figures 4, 6, 8 and 10); for brevity, these have not been reproduced here.

The EAG notes the following points regarding the hazard plots for PFS and OS:

- *PFS, cabozantinib (Figure 25).* The smoothed empirical hazard is increasing slightly over time, with a sharper increase at around 12 months, although data are limited at this time point. With the exception of the exponential model, the parametric models appear to generally reflect the

observed hazard until around 12 months. The Weibull and Gompertz models suggest an increasing hazard, but neither fully reflects the higher empirical hazard at later timepoints.

- *PFS, placebo (Figure 26).* The smoothed empirical hazard initially increases and then decreases slightly. The company's selected Weibull model does not fully reflect this shape as it has a monotonically increasing hazard in this case. The log-logistic, log-normal and generalised gamma models arguably better reflect the empirical hazard.
- *OS, cabozantinib (Figure 27).* The smoothed empirical hazard is slightly increasing over time, with a sharper increase at around 12 months. The company's selected exponential model does not reflect this shape. The generalised gamma and Gompertz models indicate an increasing hazard, but neither fully reflects the higher hazard at later timepoints.
- *OS, RPSFT-adjusted placebo (Figure 28).* The smoothed empirical hazard suggests an initially increasing then decreasing hazard. The company's selected exponential model does not reflect this shape. The log-normal, log-logistic and generalised gamma distributions feature this pattern, with the generalised gamma providing a comparatively better representation of the observed hazard.

Figure 25: Smoothed hazard versus modelled hazard for PFS, cabozantinib (reproduced from clarification response, question B2, Figure 9)

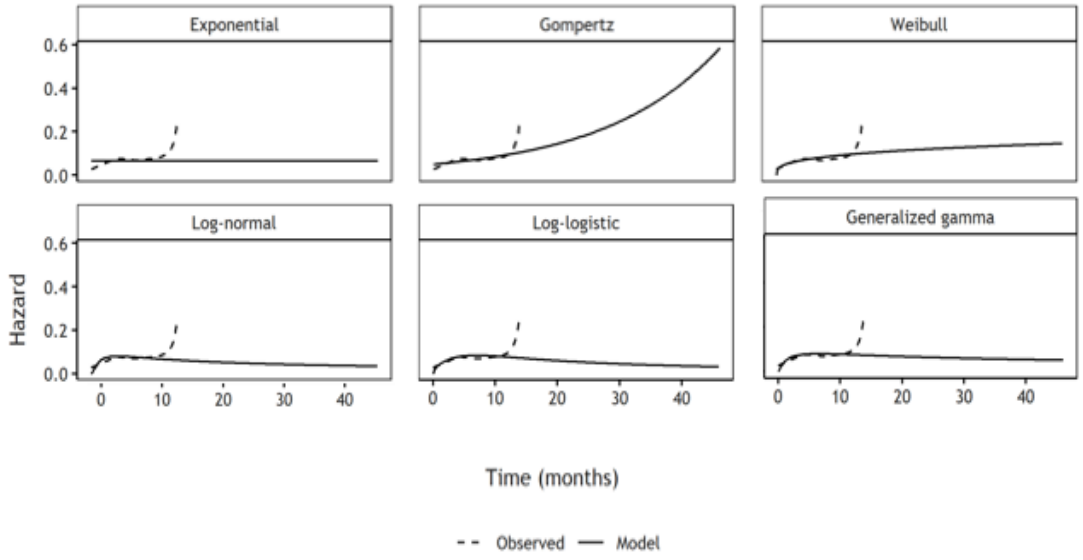


Figure 26: Smoothed hazard versus modelled hazard for PFS, placebo (reproduced from clarification response, question B2, Figure 11)

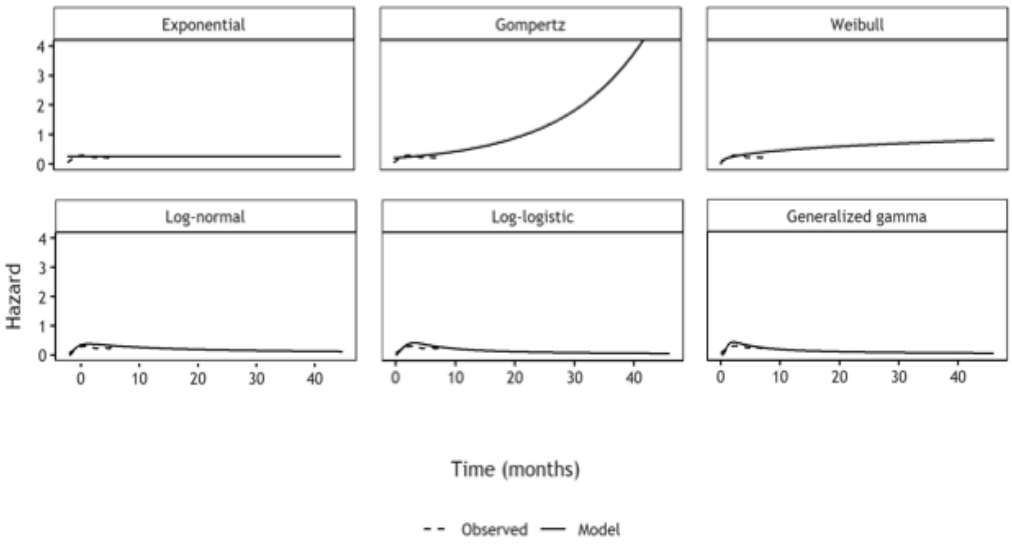


Figure 27: Smoothed hazard versus modelled hazard for OS, cabozantinib (reproduced from clarification response, question B2, Figure 5)

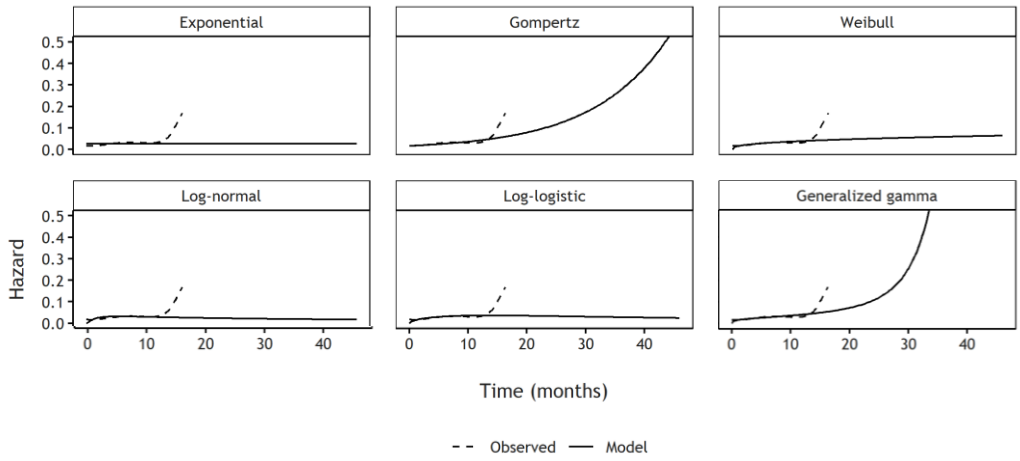
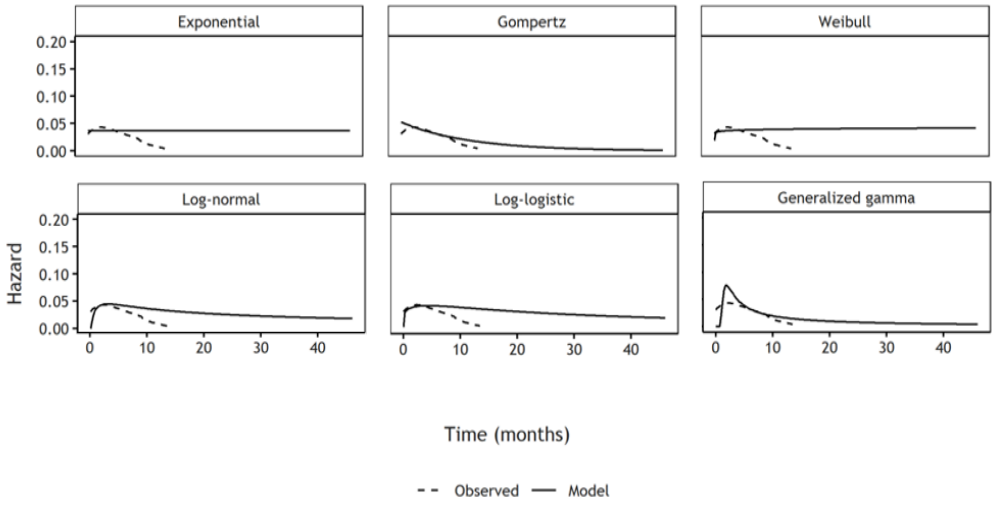


Figure 28: Smoothed hazard versus modelled hazard for OS, RPSFT-adjusted placebo (reproduced from clarification response, question B2, Figure 7)



(e) Consideration of long-term clinical plausibility

The CS[7] notes that the clinical experts who attended the company's advisory board meeting[21] raised concerns regarding the plausibility of all of the candidate parametric survival models fitted to the OS data from COSMIC-311.[19] The company's clinical advisors commented that it was not plausible that the survivor functions for OS for cabozantinib and BSC would cross – of the six models considered in the CS, only the log-normal and exponential distributions do not have this feature. The company's clinical experts also stated that ≤1% of BSC-treated patients would be expected to still be alive at 5 years. The exponential and log-normal distributions for the BSC group suggest that [redacted] and [redacted] of patients will still be alive at 5-years, respectively; therefore, these predictions are not consistent with the experts' views. The advisory board meeting minutes state that [redacted] The EAG presumes that this explains why the company has applied a structural assumption that all BSC-treated patients who are alive at 5 years die at this timepoint. However, the company's experts' concerns related to both treatment groups and no adjustment has been made to address potential over-prediction of OS in the cabozantinib group. The EAG asked the company whether their experts considered the exponential model including the 5-year death assumption to be plausible (see clarification response,[17] question B6); however, the company did not provide a response on this issue. Despite this, the EAG notes that this structural assumption results in a vertical drop in the survivor function for BSC which is not clinically realistic.

The EAG asked their own clinical advisors about their expectations of OS for patients receiving cabozantinib and BSC. Table 37 shows the EAG's clinical advisors' expectations of OS for both treatment groups at 2, 5 and 10 years, together with additional information from the minutes of the company's advisory board meeting and the company's base case model predictions.

Table 37: EAG's and company's clinical advisors' expectations of the proportions of patients alive over time for cabozantinib and BSC

Clinical advisor(s)	Cabozantinib			BSC		
	2 years	5 years	10 years	2 years	5 years	10 years
EAG Advisor 1	60-65%	30-40%	Negligible	50%	10%	Negligible
EAG Advisor 2	40-50%	20-30%	Negligible	30-40%	10-15%	Negligible
Company's clinical advisors[21]	-	-	-	-	≤1%	-
Company's model (excluding BSC 5-year death assumption)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

EAG - External Assessment Group; BSC - best supportive care

The EAG's clinical advisors commented that their estimates are uncertain. Both advisors suggested that 10-15% of BSC-treated patients would be expected to be alive at 5 years; this is higher than the

proportion suggested by the company's experts, but is consistent with the 5-year OS predicted by the company's model. One of the EAG's advisors commented that within the progressed RAI-refractory DTC population there is variation in prognosis, with some patients with more indolent disease surviving out into the longer-term without active treatment. Both of the EAG's clinical advisors expected a negligible proportion of patients on BSC to remain alive at 10 years; the company's model (excluding the 5-year death assumption) predicts that around [redacted] of patients will still be alive at this timepoint. Broadly speaking, the company's long-term model predictions of OS for BSC appear to be consistent with the EAG's advisors' expectations.

With respect to the cabozantinib group, one of the EAG's advisors suggested that 30-40% of patients might be expected to survive out to 5 years, whilst the other suggested a more pessimistic estimate of 20-30%. The company's model predicted OS at 5-years is towards the bottom of this range, at around 21%. Both of the EAG's clinical advisors suggested that a negligible proportion of patients treated with cabozantinib would be expected to remain alive at 10 years. The company's model suggests that [redacted] of patients survive out to this timepoint. This suggests that the long-term modelled OS predictions for the cabozantinib group may be overly optimistic.

Consistent with the views obtained by the company, both of the EAG's clinical advisors agreed that they would not expect the OS functions for cabozantinib and BSC to cross.

Despite the EAG's concerns regarding the fit of the exponential model to the observed data from COSMIC-311,[19] the long-term model predictions, excluding the 5-year death assumption applied in the BSC group, appear to be broadly consistent with the EAG's advisors' expectations. However, there appears to be some difference of opinion regarding expectations of long-term OS between the clinical experts consulted by the company and those consulted by the EAG.

(f) Sensitivity analysis

The CS[7] includes sensitivity analyses using the five PFS models which were not used in the base case analysis, but only one alternative OS model was considered (see Table 34). All of the scenario analyses presented in the CS use the RPSFT-adjusted OS data for BSC and all include the structural assumption that any BSC patients remaining alive at 5 years will die at this timepoint. The EAG believes that further analyses would have been useful to explore:

- Whether more flexible parametric models might better reflect the OS data and provide potentially more plausible extrapolations.
- The extent to which alternative treatment switching adjustment methods (with and without re-censoring) impact on the plausibility of the OS predictions for the BSC group.

- The impact of removing the 5-year death assumption in the BSC group.
- The impact of alternative modelling assumptions which might provide more plausible extrapolations in both groups (e.g., selecting a potentially plausible OS model for the cabozantinib group and applying an HR/AF to estimate OS for the BSC group).

EAG's conclusions regarding company's survival modelling

The EAG does not have any major concerns regarding the company's modelling of PFS. However, determining the most appropriate model for OS is more challenging for several reasons:

(a) The clinical experts consulted by the company and the EAG do not believe that it is plausible that the OS functions for cabozantinib and placebo will cross. However, the observed data from the trial suggest that the hazard for OS is decreasing in the tail of the placebo group and increasing in the tail of the cabozantinib group. Consequently, most of the company's independently-fitted models resulted in OS functions which cross and would therefore be considered clinically implausible.

(b) The company's selected exponential models for OS assume constant hazards and therefore cannot cross. However, these models do not provide a good representation of the observed OS data.

(c) The company's selected exponential models for OS implicitly assume PH, yet the empirical time-varying HR is clearly not constant, with the relative treatment effect worsening over time and favouring BSC after around 6 months.

(d) The company has not explored the use of more flexible parametric models for OS which might better reflect the observed data. The EAG agrees with the company that these models would most likely cross, but this cannot be confirmed as the company has not fitted these models.

(e) There appears to be some difference of opinion between the clinical experts consulted by the company and the EAG regarding survival expectations for patients receiving BSC alone.

Overall, the EAG believes that it is probably not possible to identify a fully parametric survival model which (i) provides a good representation of the underlying hazards for each treatment group and (ii) is clinically plausible. The EAG's exploratory analyses include the consideration of a range of alternative approaches for modelling OS (see Section 5.5).

(5) Concerns regarding health state utility values

The company's model uses health state utility values taken from the TTO valuation study reported by Fordham *et al.*[22] This study has been used in preference to the mapped EQ-5D-5L data collected in COSMIC-311.[19] The EQ-5D-5L data from COSMIC-311 are not used in the company's model, except to justify an assumption of treatment-independent utility values based on the finding that the treatment

group covariate in the mixed-effect model was not statistically significant. The EAG notes the following points:

- The EAG agrees with the company that the EQ-5D-5L data from COSMIC-311[19] are limited because data collection stopped shortly after patients progressed.
- Fordham *et al.*[22] used a TTO vignette approach valued by members of the general public. This is not in line with the NICE Reference Case.[63]
- The utility values used in the company's model are based on the adjusted regression model reported by Fordham *et al.*[22] (progression-free utility = 0.87; progressed disease utility = 0.52). As noted in critical appraisal point [1], the utility value applied in the progression-free state of the company's model is higher than the age- and sex-matched EQ-5D-3L value for the general population (utility = 0.82). This implies that it is better to have the disease than to not have the disease, which the EAG considers to be logically inconsistent.
- Utility values from Fordham *et al*[22] have been used in several previous NICE appraisals (see Table 38). However, in each of these appraisals, EQ-5D data were not collected in the clinical trials. This appraisal differs in that EQ-5D data were collected in COSMIC-311, but they have not been used in the model. These previous NICE appraisals used the observed mean utility values from Fordham *et al.*[22] (progression-free utility = 0.80, progressed disease utility = 0.50), which are lower than the utility values used in the company's model for this appraisal.
- In TA535,[12] treatment-specific utility values were used to reflect lower HRQoL for patients receiving TKIs based on EQ-5D-3L estimates sourced from the DECISION trial of sorafenib.[43]
- AE-related QALY losses are applied for 1 month. However, the company's clarification response[17] (question B21) suggests that AEs had a longer mean duration of [redacted] days (data are not reported separately by treatment group). The company's model may underestimate the negative impact of treatment-related AEs.
- One of the EAG's clinical advisors stated that overall, they would expect HRQoL for patients treated with cabozantinib to be lower than that of the general population because of toxicity. However, for patients who are fit, progression-free and are not experiencing toxicity, they would expect HRQoL to be similar to general population levels. They also agreed that the decrement associated with progression of 0.35 estimated by Fordham *et al.*[22] is plausible. The EAG's second advisor commented that whilst there may be negative psychological impacts on a patient's HRQoL due to their diagnosis of DTC, in terms of physical impacts, HRQoL would be similar to that of the general population. The advisor also commented however that patients' HRQoL whilst receiving treatment will depend on drug toxicity and that this would negatively impact on HRQoL. They also stated that the disutility value reported by Fordham *et al.* was reflective of the impact of disease progression.

Overall, the EAG believes that it may be reasonable to use the utility values reported by Fordham *et al*,[22] based on the observed means rather than the higher values obtained from the adjusted regression model, as this would provide consistency with previous NICE appraisals. It is however important to also explore the impact of: (a) alternative assumptions of AE-related disutilities on net QALY gains; (b) using the available COSMIC-311 EQ-5D-3L estimates (together with external data to inform the utility value for patients with progressed disease); and (c) applying treatment-dependent utility values which have been used in previous NICE appraisals of TKIs for DTC. These analyses are considered as part of the EAG's exploratory analyses (see Section 5.5).

Table 38: Summary of utility values used in previous NICE appraisals in advanced thyroid cancer

Appraisal	Target population for appraisal	PF utility	PD utility	Source	Elicitation/ valuation method	EAG comments
ID4046[7] (company's model)	Locally advanced/metastatic, progressed, previously treated RAI-refractory DTC	Cabozantinib/ BSC: 0.87	0.52	Fordham *et al.*[22]	TTO	No sensitivity analyses conducted around utility values
TA516[57] (AG model)	Unresectable locally advanced or metastatic MTC	Cabozantinib/ BSC: 0.80	0.50	Fordham *et al.*[22]	TTO	DECISION utility values applied in sensitivity analysis
TA535[12] (AG model)	Progressive, locally advanced or metastatic RAI-refractory DTC	Sorafenib/ lenvatinib: 0.72 BSC: 0.80	0.64	DECISION trial[43]	EQ-5D-3L	Same utility values applied in company's models
TA550[58] (company's model)	Unresectable locally advanced or metastatic MTC	Vandetanib: 0.80 BSC: 0.80	0.50	Fordham *et al.*[22]	TTO	Sensitivity analyses not detailed in committee papers
TA742[13] (company's model)	• Advanced RET fusion-positive TC who require systemic therapy • Advanced RET mutation-positive MTC who require systemic therapy	Selpercatinib/ BSC: 0.80	0.50	Fordham *et al.*[22]	TTO	DECISION utility values applied in sensitivity analysis

EAG - External Assessment Group; PF - progression-free; PD - progressed disease; TA - Technology Appraisal; RAI - radioactive iodine; TC - thyroid cancer; DTC - differentiated thyroid cancer; MTC - medullary thyroid cancer; RET - rearranged during transfection; TTO - time trade-off; EQ-5D-3L - Euroqol 5-dimensions 3-level; AG - Assessment Group

(6) Concerns regarding resource use and cost assumptions

The EAG has four concerns regarding the resource use and cost assumptions used in the company's model:

(a) *Costs associated with post-progression cabozantinib.* The company's clarification response[17] (question A15) states that the company is seeking a positive recommendation in line with the SmPC,[20] which permits treatment beyond progression if the patient is clinically benefitting from therapy and if they are not experiencing toxicity. In COSMIC-311,[19] a small proportion of patients (6.5%) received treatment beyond progression. However, the company has capped TTD by PFS, which implicitly assumes a stopping rule at progression. Given the company's intended use of cabozantinib, which includes continued treatment after progression where clinically appropriate, the EAG believes that post-progression costs should have been included in the model (i.e., the cap on TTD should be removed).

(b) *No wastage costs.* The company's model does not include any costs associated with wastage. This implies that every tablet prescribed is assumed to be taken. In reality, patients who progress or die before finishing a pack of cabozantinib will incur some drug wastage costs. The EAG believes that these costs should be accounted for in the model.

(c) *Monitoring cost assumptions*. The company's model assumes that patients receiving cabozantinib will undergo an ECG once every 6 months. One of the EAG's clinical advisors suggested that they would offer monthly ECGs during the first 3 months before moving onto 6-monthly tests, whilst the other advisor suggested a more frequent schedule of ECGs given every 2-3 months. The advisors also indicated that patients on BSC alone would be offered fewer CT scans compared with those on treatment.

(d) *TSH suppression therapy and other concomitant medication costs excluded.* The company's model does not include the costs of any concomitant therapies given as part of BSC in COSMIC-311 (e.g., indefinite TSH suppression, calcium, vitamin D, analgesics). The EAG notes that many of the concomitant therapies given in the trial are inexpensive and may not have a substantial impact on the ICER; however, these costs should have been included in the model.

(7) Weak characterisation of uncertainty

The EAG has several concerns regarding the uncertainty analysis presented in the CS.[7] These are summarised below.

(a) Some uncertain parameters held fixed in PSA.

Uncertainty surrounding the health state utility values is not modelled. Within the executable model, the cells which should contain the standard errors (SEs) around the utility values instead return "#REF!" errors. This has the effect of holding the utility values as fixed in the PSA. At the factual accuracy check

stage of the appraisal, the company clarified that they had intended to include the utility values in the PSA. Similarly, the drug administration costs for cabozantinib are uncertain parameters which are held fixed in the PSA.

(b) Additional uncertainty associated with switching adjustment is not included in the PSA.
The company's model uses the switching-adjusted time-to-event data as if they are observed (see clarification response,[17] question B16). This ignores additional uncertainty introduced through the switching adjustment, which could have been incorporated by bootstrapping the RPSFT analyses.

(c) Sampling arbitrary SEs around aggregated model functions.
For several aspects of the model, parameter uncertainty is sampled for aggregate functions of multiple parameters defined by an arbitrary SE of 10%, rather than assigning meaningful SEs to each individual underlying parameter within the function. For example, the company's model assigns a beta distribution with an SE of 10% to the total QALY loss from AEs. The EAG believes that it would be more appropriate to sample from the underlying AE frequencies and disutility values based on SEs estimated using the available data. It is unclear whether the company's approach underestimates or overestimates uncertainty.

(d) Potential reporting errors in the company's PSA in the CS.
The CS[7] reports the results of the base case analysis and scenario analyses using the deterministic and probabilistic versions of the model. Several of the ICERs reported from the probabilistic model are noticeably different from the deterministic ICERs which ordinarily would indicate non-linearity in the model. For example, the deterministic base case ICER is reported to be £27,025 whilst the probabilistic ICER is reported as £35,249. Discrepancies are also apparent between the deterministic and probabilistic scenario analyses in the CS (see deterministic analyses presented in CS, Table 48 and probabilistic analyses presented in CS, Tables 49-57). The EAG has re-run all of these analyses and did not find any major discrepancies between the probabilistic and deterministic results. The EAG believes that the company's PSA presented in the CS might have been generated using an outdated version of the model or using settings which do not reflect the company's final base case scenario. All results presented in this report are based on PSA which has been re-run by the EAG.

(e) Limited scenario analyses
The CS[7] presents scenario analyses around the discount rates, the model time horizon, age-adjustment of utility values, the PFS model and the OS model. The EAG believes that a wider set of scenarios could have been explored to provide a better assessment of decision uncertainty, including:

(a) Consideration of a wider set of models which might better represent the available OS data and/or provide more plausible OS predictions.
(b) Exploration of uncertainty around the 5-year death assumption

(c) Exploring the impact of alternative switching methods with and without re-censoring (where appropriate)
(d) The use of alternative utility sources (for example, the use of data from COSMIC-311[19] or DECISION[43]) and exploring the impact of applying treatment-dependent utility values.

These aspects of uncertainty are the focus of the EAG's exploratory analyses presented in Section 5.5.

5.4 Company's updated model provided in the company's clarification response

5.4.1 Summary of company's updated base case model

As part of their clarification response, the company submitted an revised base case model and presented the results of additional analyses undertaken using this model (see clarification response,[17] questions B5, B6, B7, B17, B19, B20, B21, B23, B24 and B27). The key features of this revised model are:

- The model errors described in critical appraisal point [1] of Section 5.3.5 have been resolved, except for issue (iii).
- OS outcomes for BSC have been amended to use an exponential model fitted to RPSFT-adjusted data from the earlier CCO1 data-cut of COSMIC-311.[19] The 5-year death assumption applied in the BSC group has been removed.
- Drug acquisition costs are adjusted using relative dose intensity (RDI) instead of compliance.
- Wastage costs are included every model cycle.

The appendix to the company's clarification response[17] provides cost-effectiveness results for a wide set of analyses using this revised model, including:

- An updated base case analysis using both the deterministic and probabilistic versions the model
- Updated one-way sensitivity analyses
- Updated scenario analyses exploring the impact of: alternative discount rates; alternative time horizons; excluding age-adjustment of utility values; using all six standard parametric models for PFS; using the Weibull model for OS; using CCO1 data for both treatment groups; reintroducing the 5-year death assumption for BSC (using CCO1 or CCO2); using the two-stage adjustment method (applied to data from CCO1 for BSC); using the company's original compliance and wastage assumptions; doubling the duration of AEs on HRQoL and using the mapped EQ-5D-3L utility estimate from COSMIC-311 in the progression-free state.

The company's revised base case results are summarised in Table 39. Compared with their original model (Table 33), the revised base case model results in a lower OS estimate for BSC, and higher incremental QALYs and costs for cabozantinib versus BSC. Excluding QALY weighting, the probabilistic version of the company's revised model suggests a base case ICER of £24,616 per QALY

gained (decision modifier = 1.2). This is lower than the probabilistic ICER of £27,169 per QALY gained generated using the company's original model.

The results of the company's updated sensitivity analyses and scenario analyses can be found in the appendix to the company's clarification response.[17] These cover a wider range of scenarios than the analyses presented in the CS.[7] For brevity, the results of these analyses are not reproduced here.

Table 39: Company's updated base case results following the clarification round (excluding QALY weighting)

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER	DM
Probabilistic model†								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£24,616**	**1.2**
BSC	[redacted]	[redacted]	[redacted]					
Deterministic model								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£24,347**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	

LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; DM - decision modifier; BSC - best supportive care
** Undiscounted*
† Based on a re-run of the PSA sub-routine by the EAG

5.4.2 EAG comments on company's updated base case model

The EAG confirms that five of the six errors discussed in Section 5.3.5 (critical appraisal point [1]) have been resolved in the company's revised model. The company's revised model adjusts drug acquisitions costs using RDI. The EAG considers that this is appropriate in most cases; however, as packs of 20mg, 40mg and 60mg cabozantinib all have the same price, the EAG believes that in this case it is more appropriate to use compliance (the proportion of days on which treatment was received) than RDI. This is because the costs of cabozantinib will be dependent on the proportion of survival time in which patients receive treatment, rather than the average dose received. The EAG also believes that the company's approach for estimating wastage costs is incorrect as the calculations assume that one quarter of a pack of cabozantinib is wasted in every cycle; the EAG believes it would be more appropriate to assume that wastage is incurred once only per patient (because they have progressed or died before completing a full pack of treatment).

The EAG understands that the company has replaced the OS data for BSC from CCO2 with those from CCO1 in an attempt to address their clinical advisors' concerns about the OS predictions in the BSC group. The updated model still uses the OS data from CCO2 in the cabozantinib group. The EAG notes the following concerns regarding this revised approach:

- It leads to inconsistent levels of follow-up between the treatment groups and excludes BSC patients who were recruited after CCO1.

- The updated model still does not provide a good fit to the placebo group data (although the company argues that plausibility is improved – see Figure 29).
- Using less data does not reduce uncertainty. Given that the CS[7] (page 48) states that the OS data from CCO2 are immature and subject to low event numbers, using data from CCO1 accentuates this problem.
- The company's original model was already broadly consistent with the EAG's clinical advisors' expectations of OS.

Overall, the EAG prefers the company's original approach to modelling OS based on CCO2, albeit with the removal of the 5-year death assumption in the BSC group. As most of the company's additional scenario analyses use CCO1 for BSC, the EAG does not consider these to be informative. However, the EAG's exploratory analyses apply most of the corrections included in the updated model (see Section 5.5).

Figure 29: Observed and modelled OS for placebo from CCO1 and CCO2 (RPSFT-adjusted)

5.5 EAG exploratory analyses

5.5.1 EAG exploratory analysis - methods

The EAG undertook exploratory analyses (EAs) using the original version of the company's model, based on CCO2 for both treatment groups. All EAs were undertaken using the deterministic version of the model. The EAG's preferred analysis was also undertaken using the probabilistic version of the

model. All analyses were undertaken by one modeller and checked by a second modeller. All analyses presented in this section reflect the PAS price of cabozantinib.

EAG's preferred analysis

The EAG's preferred analysis is comprised of five sets of amendments to the company's original model.

EA1: Correction of errors

The following corrections were applied to the company's model:

(i) *Half-cycle correction.* The company's half-cycle correction was amended such that the corrected model trace was calculated as the average of the cumulative probabilities of PFS and OS between successive cycles. TTD was not half-cycle corrected.

(ii) *Inclusion of life tables for England.* Life tables for England for the period 2018-2020 were included in the company's model.[67]

(iii) *Calculation of per-cycle mortality risks*. General population mortality risk for patients at each age was re-estimated using a weighted survival model based on the proportion of men and women recruited into COSMIC-311.[19] A structural constraint was added to the economic model to ensure that the per-cycle risk of death with the disease cannot be lower than that for the general population.

(iv) *Rounding of parameter values.* Non-rounded estimates of AE frequencies from COSMIC-311 were applied in the model.

(v) *Capped EQ-5D values.* A cap was included to ensure that HRQoL for the modelled DTC population cannot be higher than that for the general population.

(vi) *Age-adjustment of utility values.* Age-adjustment of health state utility values was implemented using the EQ-5D-3L estimates reported by Hernández Alava *et al*.,[68] based on a multiplicative approach.

These error corrections are included in all subsequent exploratory analyses.

EA2: Overall survival assumptions

Based on clinical advice received by the EAG, the 5-year mortality assumption for BSC was removed from the model. OS data from CCO2 were applied in both treatment groups, as per the company's original model.

EA3: Health utility values based on observed mean TTO estimates reported by Fordham *et al.*

Based on clinical advice received by the EAG, and for purposes of consistency with the majority of previous NICE appraisals of treatments for TC, the health state utility values were amended to reflect the observed mean TTO values reported by Fordham *et al.*[22] The utility values applied in the progression-free and progressed disease states were 0.80 and 0.50, respectively.

EA4: Stopping rule removed, TTD modelled using Weibull distribution

In order to align modelled costs and outcomes with the experience of the COSMIC-311 trial,[19] and for consistency with the company's intended use of cabozantinib, the cap which constrains TTD by PFS was removed from the model. The TTD function was modelled using a Weibull distribution - this is the second-best fitting TTD model and this distribution produces a smaller gap between PFS and TTD compared with the exponential TTD model used in the company's base case.

EA5: Drug wastage costs

The total drug cost calculations were amended to include 7 days' wastage of cabozantinib per patient over their lifetime.

EA6: EAG's preferred analysis

The EAG's preferred analysis includes EA1-EA5 inclusive.

Additional sensitivity analyses

Five sets of additional sensitivity analyses (ASAs) were undertaken using the EAG's preferred model (EA6).

ASA1: Alternative OS assumptions

Three alternative approaches were used to explore the uncertainty around modelled OS gains for cabozantinib versus BSC:

ASA1a - Treatment effect waning at 3 years. This model retains the exponential models applied in the EAG's preferred analysis, but applies the OS hazard for BSC to both groups after 36 months.

ASA1b - Hybrid Kaplan-Meier up to 12 months plus exponential tail with constant HR. This analysis uses the Kaplan-Meier estimates for both treatment groups up to 12 months (note the maximum follow-up in the re-censored RPSFT-adjusted group is only slightly longer, at [redacted] months). After 12 months, OS is modelled using the exponential models applied in the company's base case. This analysis assumes a constant HR on OS for cabozantinib versus BSC.

ASA1c - Hybrid Kaplan-Meier up to 12 months plus exponential tail based on BSC group hazard. This analysis is the same as ASA1b, except that after 12 months, the hazard rate for BSC is applied in both treatment groups. This analysis therefore assumes an HR equal to 1.0 after 12 months.

The OS assumptions employed in the EAG's preferred analysis and ASA1a-c are shown graphically in Figure 30 and Figure 31.

Figure 30: OS models used in EA6 (no treatment effect waning) and ASA1a (treatment effect waning at 3 years)

OS - overall survival; EA - exploratory analysis; ASA - additional sensitivity analysis; BSC - best supportive care
Note: OS for BSC is the same for EA6 and ASA1a

Figure 31: OS models used in ASA1b (hybrid model, constant HR) and ASA1c (hybrid model, HR=1.0)

OS - overall survival; ASA - additional sensitivity analysis; BSC - best supportive care; HR - hazard ratio
Note: OS for BSC is the same for ASA1b and ASA1c

ASA2: Progression-free utility value based on COSMIC-311

Two sensitivity analyses were undertaken to explore the impact of using alternative sources of health state utility values:

- *ASA2a - COSMIC-311 utility value applied in progression-free state.* The utility value for the progression-free state in both groups was amended to reflect the mean utility value for patients without progression in COSMIC-311[19] (utility value = [redacted]). The utility value for the progressed disease state was not amended.
- *ASA2b – DECISION trial utility values (treatment-dependent)*. The utility values from the DECISION trial[43] were applied (progression-free: TKI utility = 0.72; BSC utility = 0.80; progressed disease [both groups]: utility = 0.64). QALY losses associated with AEs were removed from the model as their inclusion would likely double-count the impact of toxicity. The EAG notes that this scenario is pessimistic in that it assumes that toxicity impacts associated with cabozantinib persist for the entire duration in which patients remain progression-free.

ASA3: AE decrement doubled

Within this analysis, QALY losses associated with AEs for cabozantinib and BSC were assumed to persist for 2 months.

ASA4: ECG costs doubled

Within this analysis, the frequency of ECGs required for patients receiving cabozantinib was doubled.

ASA5: CT scans excluded for BSC

Within this analysis, patients receiving BSC were assumed not to undergo any CT scans.

5.5.2 EAG exploratory analysis – results (excluding QALY weighting)

The results of the EAG's preferred analysis are presented in Table 40. The EAG's analyses indicate that the correction of errors increases the company's deterministic base case ICER from £28,148 to £29,080 per QALY gained (EA1). Removing the 5-year death assumption increases the EAG's error-corrected ICER to £32,747 per QALY gained (EA2). Removing the stopping rule increases the EAG's error-corrected ICER to £33,050 per QALY gained (EA4). Using the observed mean values from Fordham *et al.*[22] and including drug wastage costs have only a small impact on the ICER (EA3 and EA5). The EAG's preferred ICER based on the probabilistic model is estimated to be £39,347 per QALY gained (EA6b); the deterministic version of the model suggests a lower ICER of £38,876 per QALY gained (EA6a). The decision modifier is 1.2 across all analyses. The results of the EAG's preferred analysis should be interpreted with some caution due to the EAG's concerns regarding the poor fit of the OS models and the implicit assumption of a constant treatment effect on OS for cabozantinib.

Table 40: EAG preferred analysis results (excluding QALY weighting)

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER	DM
Company's base case (deterministic)								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£28,148**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA1: Correction of errors								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£29,080**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA2: Remove 5-year death assumption for BSC								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£32,747**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA3: Observed mean utility values from Fordham *et al.*								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£29,834**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA4: Stopping rule removed, TTD modelled using Weibull distribution								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£33,050**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA5: Inclusion of drug wastage costs								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£29,623**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA6a: EAG preferred analysis (deterministic)								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£38,876**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA6b: EAG preferred analysis (probabilistic)								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£39,347**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	

LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; DM - decision modifier; EA - exploratory analysis; RDI - relative dose intensity
** Undiscounted*

Table 41 presents the results of the EAG's additional sensitivity analyses. The inclusion of an assumption of treatment effect waning at 3 years increases the EAG's preferred deterministic ICER from £38,876 to £47,987 per QALY gained (ASA1a). The use of a hybrid model has a limited impact on the ICER if the constant HR is retained after 12 months (ASA1b; ICER = £40,675 per QALY gained), but a substantial impact if an HR of 1.0 is assumed (ASA1c; ICER = £71,087 per QALY gained). Applying alternative utility values from COSMIC-311 or DECISION increases the ICER to around £44,000 per QALY gained (ASA2a and ASA2b). Alternative assumptions regarding the duration of AEs and the frequency of ECGs and CT scans have only a limited impact, with each analysis increasing the ICER by less than £1,000 (ASA3, ASA4 and ASA5).

Table 41: EAG additional sensitivity analyses results (excluding QALY weighting)

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER	DM
EA6: EAG preferred analysis (deterministic)								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£38,876**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA1a: Exponential OS with treatment effect waning at 3 years								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£47,987**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA1b: Hybrid KM + exponential tail after 12 months, constant HR								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£40,675**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA1c: Hybrid KM + exponential tail after 12 months, BSC hazard rate in both groups								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£71,087**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA2a: COSMIC-311 utility value in progression-free state								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£44,833**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA2b: DECISION trial utility values								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£44,302**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA3: AE QALY losses doubled								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£39,395**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA4: ECG costs doubled								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£39,461**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA5: CT scan costs for BSC removed								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£39,200**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	

LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; DM - decision modifier; EA - exploratory analysis; RDI - relative dose intensity
** Undiscounted*

5.5.3 EAG exploratory analysis – results (including QALY weighting)

Table 42 summarises the results of the company's base case analysis, the EAG's preferred analysis and the EAG's additional sensitivity analyses with and without QALY weighting using a decision modifier of 1.2. For each analysis, results are presented in terms of ICERs and incremental net health benefits (NHBs) assuming WTP thresholds of £20,000 and £30,000 per QALY gained. When QALY weighting is included, the ICER is above £30,000 per QALY gained for all analyses (i.e., the incremetnal NHB is negative), except for the company's base case.

Table 42: Results of EAG's preferred analysis

Analysis	Excluding QALY weighting			Including QALY weighting		
	ICER	Inc. NHB (λ=£20k)	Inc. NHB (λ=£30k)	ICER	Inc. NHB (λ=£20k)	Inc. NHB (λ=£30k)
Company's original base case (deterministic)	£28,148	[redacted]	[redacted]	£23,456	[redacted]	[redacted]
Company's original base case (probabilistic)	£27,169	[redacted]	[redacted]	£22,641	[redacted]	[redacted]
EA6: EAG preferred analysis (deterministic)	£38,876	[redacted]	[redacted]	£32,397	[redacted]	[redacted]
EA6: EAG preferred analysis (probabilistic)	£39,347	[redacted]	[redacted]	£32,789	[redacted]	[redacted]
ASA1a: Exponential OS with treatment effect waning at 3 years	£47,987	[redacted]	[redacted]	£39,989	[redacted]	[redacted]
ASA1b: Hybrid KM + exponential tail after 12 months, constant HR	£40,675	[redacted]	[redacted]	£33,895	[redacted]	[redacted]
ASA1c: Hybrid KM + exponential tail after 12 months, BSC hazard rate in both groups	£71,087	[redacted]	[redacted]	£59,240	[redacted]	[redacted]
ASA2a: COSMIC-311 utility value in progression-free state	£44,833	[redacted]	[redacted]	£37,361	[redacted]	[redacted]
ASA2b: DECISION trial utility values	£44,302	[redacted]	[redacted]	£36,918	[redacted]	[redacted]
ASA3: AE QALY losses doubled	£39,395	[redacted]	[redacted]	£32,829	[redacted]	[redacted]
ASA4: ECG costs doubled	£39,461	[redacted]	[redacted]	£32,884	[redacted]	[redacted]
ASA5: CT scan costs for BSC removed	£39,200	[redacted]	[redacted]	£32,667	[redacted]	[redacted]

QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; NHB - net health benefit; EA - exploratory analysis; ASA - additional sensitivity analysis; KM - Kaplan-Meier; HR - hazard ratio; BSC - best supportive care; AE - adverse event; ECG - electrocardiogram; CT - computerised tomography
Note - all analyses are based on the determistic version of the model unless otherwise stated

5.6 Discussion

The CS[7] includes an SLR of existing economic studies of treatments for RAI-refractory DTC and reports the methods and results of a *de novo* model-based health economic analysis of cabozantinib for RAI-refractory DTC.

The company's SLR identified five published economic analyses and one clinical modelling study. All but one of these analyses were available only in abstract form and none assessed the cost-effectiveness of cabozantinib for RAI-refractory DTC.

The company's economic model assesses the cost-effectiveness of cabozantinib versus BSC alone for locally advanced or metastatic, progressed, RAI-refractory DTC. The model adopts a partitioned survival approach which includes three health states: (i) progression-free; (ii) progressed disease and (iii) dead. The analysis adopts an NHS and PSS perspective, including QALYs accrued by DTC

patients; caregiver effects are not included. Clinical outcomes for both groups are based on parametric survival models fitted to data on PFS and OS from COSMIC-311,[19] including adjustment of OS in the BSC group to account for treatment switching in the placebo arm of the trial. The company's base case analysis assumes that cabozantinib would be discontinued at progression, although the company is seeking a positive recommendation which includes treatment beyond progression in patients who are still deriving clinical benefit from cabozantinib. Health state utility values are based on data from an external TTO study;[22] resource use and cost parameters were based on COSMIC-311, clinical input obtained as part of previous NICE TAs,[12, 13] previous literature[53] and standard costing sources.[51, 52]

The probabilistic version of the company's original model suggests that the ICER for cabozantinib versus BSC is £27,169 per QALY gained (excluding QALY weighting). The deterministic ICER is similar (£28,148 per QALY gained). Based on the discounted QALYs predicted for the comparator group, the absolute shortfall is estimated to be 9.62 QALYs, whereas the proportional shortfall is estimated to be 89%, thereby leading to a decision modifier of 1.2. Including QALY weighting in the calculations leads to a probabilistic ICER of £22,641 per QALY gained.

The EAG critically appraised the company's health economic analysis and double-programmed the deterministic version of the company's original model. The EAG's main concerns regarding the company's submitted economic model relate to the approach used to model OS:

(a) The company's original economic model uses exponential distributions to model OS together with a structural assumption that all BSC-treated patients who are still alive at 5 years will die at this timepoint. The exponential models do not provide a good representation of the available OS data, particularly for the BSC group, and the structural assumption applied in the BSC group forces a vertical drop in OS which is not clinically realistic.

(b) The OS data at CCO2 are immature. However, the available data suggest that the hazard of death is increasing at the tail in the cabozantinib group and slowing at the tail in the BSC group, indicating that the treatment effect for cabozantinib over BSC is worsening over time. The use of exponential models, regardless of whether they are independently or jointly fitted, implies a constant HR for OS which does not reflect what has been observed in the trial.

(c) Most of the standard parametric survival models lead to the OS functions for cabozantinib and BSC crossing, which the company's and EAG's clinical experts do not consider to be plausible. The company has not explored more flexible models (e.g., RCS models); however, it is likely that these would also lead to the OS functions crossing.

(d) There is some difference of opinion amongst the clinical experts regarding expectations of long-term OS for cabozantinib and BSC. Clinical input obtained by the company suggests that the exponential distribution (excluding the 5-year BSC group death assumption) overestimates OS in both treatment groups. The EAG's clinical advisors' expectations of OS for BSC were consistent with the company's exponential model; however, OS for the cabozantinib group may be overestimated.

The EAG also identified several other issues which impact on the ICER, including minor programming errors, uncertainty around the most appropriate source(s) to use to inform health state utility values, and the exclusion of the costs of post-progression cabozantinib and drug wastage.

The EAG's preferred model includes: (i) the correction of errors; (ii) removal of the 5-year death assumption for BSC; (iii) the use of observed mean utility values reported by Fordham *et al.*;[22] (iv) the inclusion of post-progression cabozantinib costs and (v) the inclusion of drug wastage costs. Excluding QALY weighting, the EAG's preferred analysis using the probabilistic version of the model (EA6b) suggests that the ICER for cabozantinib versus BSC expected to be £39,347 per QALY gained. When QALY weighting is included, the probabilistic ICER is £32,789 per QALY gained. However, this analysis should be interpreted with caution as the model does not provide a good fit to the available OS data from COSMIC-311[19] and it assumes a constant treatment effect on OS which contrasts with what has been observed in the trial. This scenario is therefore likely to be optimistic. The EAG explored three alternative approaches to modelling OS, each of which suggest that the ICER for cabozantinib may be higher than the EAG's preferred analysis. However, each of these approaches also make strong assumptions regarding long-term OS benefits for cabozantinib:

- *ASA1a. EAG-preferred model + treatment effect waning at 3 years.* ICER excluding QALY weighting = £47,987; ICER including QALY weighting = £39,989. This model produces less optimistic predictions of OS which are more consistent with the expectations of the EAG's clinical advisors; however, the timepoint at which waning is applied is arbitrary.
- *ASA1b. Hybrid Kaplan-Meier plus exponential tail after 12 months.* ICER excluding QALY weighting = £40,675; ICER including QALY weighting = £33,895. This model reflects the observed data from COSMIC-311 up to 12 months, but retains a constant HR on OS thereafter. This is optimistic given the trial data.
- *ASA1c. Hybrid Kaplan-Meier plus BSC group exponential tail after 12 months.* ICER excluding QALY weighting = £71,087; ICER including QALY weighting = £59,240. This model reflects the observed data from COSMIC-311 up to 12 months, but assumes an HR of 1.0 thereafter. This may be pessimistic as it assumes virtually no incremental OS benefit after 12 months.

Longer-term follow-up would help to reduce uncertainty around the relative treatment effect of cabozantinib on OS. However, the company has confirmed that there are no further planned beyond CCO2 (see clarification response,[17] question A20).

The EAG's exploratory analyses suggest that applying alternative utility values and cost assumptions increase the ICER for cabozantinib; however, these factors are likely to have a smaller impact on the ICER compared with the assumptions about OS.

6. OVERALL CONCLUSIONS

Clinical effectiveness conclusions

The primary clinical efficacy and safety evidence for cabozantinib in adult patients with RAI-refractory DTC receiving second-line or third-line treatment detailed in the CS comes from COSMIC-311. This was an international Phase III, multi-centre, placebo-controlled, blinded RCT which assessed cabozantinib at the licensed dose of 60mg daily. The EAG's clinical advisors considered that the trial population is similar to the population who would be eligible to receive cabozantinib in UK clinical practice. The trial had two clinical cut-offs: the primary clinical cut-off date was the 19th of August 2020 (number of patients: 125 cabozantinib; 62 placebo) (CCO1); and, after further enrolment, the second, 'supportive analyses' data cut-off date was the 8th of February 2021 (170 cabozantinib; 88 placebo) (CCO2).

Cabozantinib demonstrated significant efficacy compared with placebo in terms of PFS and ORR at both data cut-offs. However, COSMIC-311 was a medium-sized trial with 258 subjects across two arms at CCO2, but with only a short length of follow-up (median 10.1 months at the latest data cut-off, CCO2, and 6.2 months at the primary data cut-off, CCO1). The trial was assessed by the EAG as being at high risk of bias on account of the deviation from the pre-specified interventions: sizeable proportions of patients with progressive disease in the placebo arm crossed-over to receive open-label treatment with cabozantinib therapy within a median period of only 1.9 months after commencement of the trial (31% at CCO1 and 45% at CCO2), which potentially confounded the outcomes of OS and safety. The CS acknowledges that there was no significant difference between the two arms in terms of OS, only a trend favouring cabozantinib, even after adjusting for treatment switching. No meta-analysis was conducted despite the presence of a single-arm trial that satisfied the criteria (NCT01811212). ITCs were not undertaken because of the absence of comparable trials of second-line therapy in the relevant population and due to the availability of direct evidence from COSMIC-311.

There were high rates of treatment-related AEs and SAEs in the cabozantinib arm compared with the placebo arm, as well as dose modifications due to AEs. A number of AEs related to cabozantinib were frequent: diarrhoea, PPES, hypertension, fatigue, hypocalcaemia, [redacted], and decreased appetite. The incidence of these individual AEs at Grade 3 or higher was however generally low (<12%). HRQoL was only assessed by EQ-5D in patients who had not progressed / up to the point of progression (to prevent confounding due to crossover) and no significant or clinically important difference between cabozantinib and placebo was found for patients who had not progressed up to 33 weeks (there were only five or fewer patients in the placebo arm after this point, preventing a meaningful comparison from being made).

Cost-effectiveness conclusions

The company's model provides an economic comparison of cabozantinib versus BSC for adult patients with locally advanced or metastatic DTC, refractory or not eligible to RAI who have progressed during or after prior systemic therapy. Excluding QALY weighting, the probabilistic version of the company's original model suggests a base case ICER of £27,169 per QALY gained. Based on a decision modifier of 1.2, the ICER including QALY weighting is estimated to be £22,641 per QALY gained.

The EAG's preferred analysis includes: (i) the correction of errors; (ii) the removal of the 5-year death assumption for BSC; (iii) the use of observed mean utility values reported by Fordham *et al.*; (iv) the inclusion of post-progression cabozantinib costs and (v) the inclusion of drug wastage costs. Excluding QALY weighting, the EAG's preferred analysis suggests that the probabilistic ICER for cabozantinib versus BSC is £39,347 per QALY gained. When QALY weighting is included, the probabilistic ICER is £32,789 per QALY gained.

The EAG notes that there is considerable uncertainty around the modelled estimates of OS. Additional sensitivity analyses undertaken by the EAG suggest that less optimistic assumptions regarding OS may lead to higher ICERs which are in excess of £71,000 per QALY gained when QALY weighting is excluded, and in excess of £59,000 per QALY gained when QALY weighting is included. The inclusion of alternative sources of health state utility values and/or cost assumptions also increase the ICER, but may have a comparatively smaller impact than the OS assumptions.

7. REFERENCES

1. Cancer Research UK. Thyroid cancer statistics. CRUK: London, UK; 2022.
2. Kent WD, Hall SF, Isotalo PA, Houlden RL, George RL, Groome PA. Increased incidence of differentiated thyroid carcinoma and detection of subclinical disease. *Canadian Medical Association Journal* 2007;177:1357-61.
3. Tumino D, Frasca F, Newbold K. Updates on the management of advanced, metastatic, and radioiodine refractory differentiated thyroid cancer. *Frontiers in Endocrinology* 2017;8.
4. Hundahl SA, Fleming ID, Fremgen AM, Menck HR. A national cancer data base report on 53,856 cases of thyroid carcinoma treated in the U.S., 1985-1995. *Cancer* 1998;83:2638-48.
5. Jonklaas J, Sarlis NJ, Litofsky D, Ain KB, Bigos ST, Brierley JD*, et al.* Outcomes of patients with differentiated thyroid carcinoma following initial therapy. *Thyroid* 2006;16:1229-42.
6. Fleeman N, Houten R, Bagust A, Richardson M, Beale S, Boland A*, et al.* Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine: a systematic review and economic evaluation. *Health Technology Assessment* 2017;24:1-229.
7. Ipsen. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Document B - company's evidence submission. Ipsen: Slough, UK; 2022.
8. Lirov R, Worden FP, Cohen MS. The treatment of advanced thyroid cancer in the age of novel targeted therapies. *Drugs* 2017;77:733-45.
9. Durante C, Haddy N, Baudin E, Leboulleux S, Hartl D, Travagli JP*, et al.* Long-term outcome of 444 patients with distant metastases from papillary and follicular thyroid carcinoma: benefits and limits of radioiodine therapy. *Journal of Clinical Endocrinology and Metabolism* 2006;91:2892-9.
10. Oh HS, Shin DY, Kim M, Park SY, Kim TH, Kim BH*, et al.* Extended real-world observation of patients treated with sorafenib for radioactive iodine-refractory differentiated thyroid carcinoma and impact of lenvatinib salvage treatment: A Korean multicenter study. *Thyroid* 2019;29:1804-10.
11. Brilli L, Dalmiglio C, Pilli T, Barbato F, Maino F, Capezzone M*, et al.* Improvement of overall survival using TKIs as salvage therapy in advanced thyroid carcinoma: Real-Life data on a single center experience. *Journal of Clinical Medicine* 2021;10:384.
12. National Institute for Health and Care Excellence. TA535 - Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine. NICE: London, UK; 2018.
13. National Institute for Health and Care Excellence. TA742 - Selpercatinib for treating advanced thyroid cancer with RET alterations. NICE: London, UK; 2021.
14. National Institute for Health and Care Excellence. TA630 - Larotrectinib for treating NTRK fusion-positive solid tumours. NICE: London, UK; 2020.
15. National Institute for Health and Care Excellence. TA644 - Entrectinib for treating NTRK fusion-positive solid tumours. NICE: London, UK; 2020.
16. Filetti S, Durante C, Hartl D. ESMO Clinical Practice Guideline update on the use of systemic therapy in advanced thyroid cancer. *Annals of Oncology* 2022;33:674-84.
17. Ipsen. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Company's response to clarification questions from the EAG. Ipsen: Slough, UK; 2022.
18. National Institute for Health and Care Excellence. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Final scope. NICE: London, UK; 2022.

19. Brose MS, Robinson B, Sherman SI, Krajewska J, Lin CC, Vaisman F*, et al.* Cabozantinib for radioiodine-refractory differentiated thyroid cancer (COSMIC-311): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncology* 2021;22:1125-38.

20. European Medicines Agency. Summary of product characteristics - cabometyx. EMA: London, UK; 2022.

21. Ipsen. IPSEN HTA advisory board for differentiated thyroid cancer insights report. Ipsen: Slough, UK; 2022.

22. Fordham BA, Kerr C, Freitas HM, Lloyd AJ, Johnston K, Pelletier CL. Health state utility valuation in radioactive iodine-refractory differentiated thyroid cancer. *Patient Preference and Adherence* 2015;9:1561-72.

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

24. Suteau V, Munier M, Briet C, Rodien P. Sex Bias in Differentiated Thyroid Cancer. *International Journal of Molecular Science* 2021;22.

25. Ipsen. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Company's submission appendices. Ipsen: Slough, UK; 2022.

26. Brose MS, Robinson B, Sherman SI, Jarzab B, Lin CC, Vaisman F. Cabozantinib versus placebo in patients with radioiodine-refractory differentiated thyroid cancer who have progressed after prior VEGFR-targeted therapy: Results from the phase 3 COSMIC-311 trial. *J Clin Oncol* 2021;May 20;39(15_suppl):6001–6001.

27. Capdevila J, Sherman B, Jarzab S, Lin B, Vaisman C, Hoff F*, et al.* LBA67 Cabozantinib versus placebo in patients with radioiodine-refractory differentiated thyroid cancer who have progressed after prior VEGFR-targeted therapy: Updated results from the phase III COSMIC-311 trial and prespecified subgroup analyses by prior therapy. *Annals of Oncology* 2021;32:S1343.

28. Robinson A, et al. Effect of age on efficacy and safety of cabozantinib versus placebo in patients with radioiodine (RAI)-refractory differentiated thyroid cancer (DTC) with progression after VEGFR-targeted therapy: subgroup analysis from the phase 3 COSMIC 311 study. Poster presented at: 90th Annual Meeting of the American Thyroid Association; 2021. 2022.

29. Durante C, Robinson B, Sherman S, Krajewska J, Lin C, Vaisman F*, et al.* Effect of age on efficacy and safety of cabozantinib versus placebo in patients with radioiodine (RAI)-refractory differentiated thyroid cancer (DTC) with progression after VEGFR-targeted therapy: subgroup analysis from the phase 3 COSMIC 311 Study. *Endocrine Abstracts* 2022;81:OC3.4.

30. Durante C, Robinson B, Sherman S, Krajewska J, Lin C, Vaisman F*, et al.* Cabozantinib versus placebo in patients with radioiodine-refractory differentiated thyroid cancer (DTC) who have progressed after prior VEGFR-Targeted therapy: updated results from the phase 3 COSMIC-311 trial and prespecified subgroup analyses based on prior therapy. *Endocrine Abstracts* 2022;81:OC3.2.

31. Hernando J, Capdevila J, Robinson B, Sherman S, Jarząb B, Lin CC*, et al.* Cabozantinib (C) versus placebo (P) in patients (pts) with radioiodine-refractory (RAIR) differentiated thyroid cancer (DTC) who have progressed after prior VEGFR-targeted therapy: Outcomes in prespecified subgroups based on prior VEGFR-targeted therapy. *Journal of Clinical Oncology* 2022;40.

32. Capdevila J, Robinson B, Sherman S, Jarząb B, Lin C, Vaisman F*, et al.* Cabozantinib versus placebo in patients (pts) with radioiodine-refractory (RAIR) differentiated thyroid cancer (DTC) who progressed after prior VEGFR-targeted therapy: Outcomes in prespecified subgroups based on histology subtypes. *Journal of Clinical Oncology* 2022;40:6081.

33. Brose MS, Shenoy S, Bhat N, Harlacker AK, Yurtal RK, Posey ZA. A phase II trial of cabozantinib (CABO) for the treatment of radioiodine (RAI)-refractory differentiated thyroid carcinoma (DTC) in the first-line setting. *Journal of Clinical Oncology* 2018;May 20;36(15_suppl):6088–6088.
34. Cabanillas MEDS, Geyer JA, Wirth S, Menefee LJ, Liu ME, Shah SV. Cabozantinib as salvage therapy for patients with tyrosine kinase inhibitor-refractory differentiated thyroid cancer: Results of a multicenter phase II international thyroid oncology group trial. *Journal of Clinical Oncology* 2017;35:3315-21.
35. Centre for Reviews and Dissemination. Systematic reviews - CRD's guidance for undertaking reviews in health care. CRD: York, UK; 2009.
36. Suzuki C, Kiyota N, Imamura Y, Goto H, Suto H, Chayahara N*, et al.* Exploratory analysis of prognostic factors for lenvatinib in radioiodine-refractory differentiated thyroid cancer. *Head and Neck* 2019;41:3023-32.
37. Manohar P, Beesley L, Bellile E, Worden F, Avram A. Prognostic value of FDG-PET/CT metabolic parameters in metastatic radioiodine-refractory differentiated thyroid cancer. *Clinical Nuclear Medicine* 2018;43:641-7.
38. Sterne J, Savović J, Page M, Elbers R, Blencowe N, Boutron I*, et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *British Medical Journal* 2019;366:14898.
39. Exelixis Inc. XL184-311 Clinical study report (COSMIC-311, CCO1). 2020.
40. Exelixis Inc. XL184-311 Clinical study report - Addendum 1 (COSMIC-311, CCO2); 2021.
41. Exelixis Inc. XL184-311 Clinical study report - Addendum 2 (COSMIC-311, CCO2); 2021.
42. Schlumberger M, Tahara M, Wirth LJ, Robinson B, Brose MS, Elisei R*, et al.* Lenvatinib versus placebo in radioiodine-refractory thyroid cancer. *New England Journal of Medicine* 2015;372:621-30.
43. Brose MS, Nutting CM, Jarzab B, Elisei R, Siena S, Bastholt L. Sorafenib in locally advanced or metastatic, radioactive iodine-refractory, differentiated thyroid cancer: a randomized, double-blind, phase 3 trial. *Lancet* 2014;384:319-28.
44. Wilson L, Huang W, Chen L, Ting J, Cao V. Cost effectiveness of lenvatinib, sorafenib and placebo in treatment of radioiodine-refractory differentiated thyroid cancer. *Thyroid* 2017;27:1043-52.
45. Carlson J, Suh K, Xia F, Williamson T, Sullivan S. PCN11 The potential long-term comparative effectiveness of larotrectinib vs. lenvatinib or sorafenib for treatment of NTRK fusion-positive metastatic thyroid cancer. *Value in Health* 2021;24:S20.
46. Erdal E, Sayman H, Turkmen C, Aral F, Yildiz O, Okutur K. Cost-effectiveness of sorafenib for treatment of radioactive iodine (RAI)-refractory locally advanced/metastatic differentiated thyroid cancer (DTC) in Turkey. *Value in Health* 2015;18:PA203-A4.
47. Tremblay G, Pelletier C, Copher R, Forsythe A, Majethia U. Cost-effectiveness analysis of lenvatinib as a treatment for radioactive iodine refractory differentiated thyroid cancer in the United States. *Value in Health* 2016;19:PA151.
48. Huang W, Chen L, Ting J, Cao V, Sung H, Yokokura M. The EVPI of treatment strategies for radioiodine-refractory differentiated thyroid cancer. *Value in Health* 2016;19:PA886.
49. Carrasquilla-Sotomayor M, Alvis-Zakzuk N, Rosa F, Alvis Zakzuk J, Marrugo Figueroa R, Miranda Machado P. Cost and effectiveness of the use of sorafenib in differentiated thyroid cancer. *Value in Health* 2017;20:PA876.
50. Joint Formulary Committee. British National Formulary. BMJ Group and Pharmaceutical Press: London, UK; 2022.
51. NHS England. National cost collection: National schedule of NHS costs 2020-21. NHSE: London, UK: NHSE; 2022.

52. Jones K, Burns A. Unit costs of health and social care 2021. Personal Social Services Research Unit: Kent, UK; 2021.
53. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. The Nuffield Trust: London, UK; 2014.
54. Brostrom G. Package 'eha' - R package. Available from https://cran.r-project.org/web/packages/eha/eha.pdf; 2017.
55. Allison A, White IR, Bond S. rpsftm: An R package for rank preserving structural failure time models. *The R Journal* 2017;9:342-53.
56. Hernández-Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. *Journal of Health Economics* 2017;55:139-52.
57. National Institute for Health and Care Excellence. Technology Appraisal 516 - Cabozantinib for treating medullary thyroid cancer. NICE: London, UK; 2018.
58. National Institute for Health and Care Excellence. TA550 - Vandetanib for treating medullary thyroid cancer. NICE: London; 2018.
59. Gallop K, Kerr C, Simmons S, McIver B, Cohen EE. A qualitative evaluation of the validity of published health utilities and generic health utility measures for capturing health-related quality of life (HRQL) impact of differentiated thyroid cancer (DTC) at different treatment phases. *Quality of Life Research* 2015;24:325-38.
60. McNamara S, Schneider PP, Love-Koh J, Doran T, Gutacker N. Quality-adjusted life expectancy norms for the English population. *Value in Health* 2022;Epub ahead of print.
61. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th edn. New York, USA: Oxford University Press; 2015.
62. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C*, et al.* Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value in Health* 2003;6:9-17.
63. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. NICE: London, UK; 2022.
64. National Institute for Health and Care Excellence. TA498 - Lenvatinib with everolimus for previously treated advanced renal cell carcinoma. NICE: London, UK; 2018.
65. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. DSU: Sheffield, UK; 2013.
66. Rutherford MJ, Lambert PC, Sweeting MJ, Pennington B, Crowther MJ, Abrams KR*, et al.* NICE DSU TSD 21: Flexible methods for survival analysis. DSU: Sheffield, UK; 2020.
67. Office for National Statistics. National life tables: England 2018-2020. ONS: London, UK; 2021.
68. Hernández-Alava M, Pudney S, Wailoo A. Estimating EQ-5D-3L by age and sex for the UK. DSU: Sheffield, UK; 2022.

Single Technology Appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on 2 December 2022** using the below 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.

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

Issue 1 Missing reference

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 13: Missing reference for Fordham (2015) "Health state utility values are based on estimates reported from an external TTO valuation study of RAI-refractory DTC health states (Fordham et al.)"	- Fordham BA, Kerr C, Freitas HM, Lloyd AJ, Johnston K, Pelletier CL. Health state utility valuation in radioactive iodine-refractory differentiated thyroid cancer. Patient Preference and Adherence 2015;9:1561-72.	Minor adjustment for clarity	This is not a factual inaccuracy. We have not included any references in the executive summary.

Issue 2 Factual inaccuracy

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 17: Request to include the revised model base case for reference "Company's base case model"	Appendix of EAG clarification questions includes all results	Major adjustment for transparency	This is not a factual inaccuracy – as mentioned on page 14, we do not consider the revised model to be suitable for decision-making. However, for the sake of transparency, we

			have included the company's revised deterministic base case results in Table 2 of the executive summary.

Issue 3 Factual inaccuracy

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 21 (Section 2.2.3). "One of the clinical advisors stated that they would not offer cabozantinib to patients who have a prolonged QT interval, but highlighted that this reflects a minority of patients." The reader of the report who will have noted lenvatinib and sorafenib are available a first-line treatment options and one of which will have been used prior to treatment with cabozantinib may possibly	Suggest either clarifying this statement with the expert or adding an additional sentence as follows: "One of the clinical advisors stated that they would not offer cabozantinib to patients who have a prolonged QT interval, but highlighted that this reflects a minority of patients. The Summary of Product Characteristics (SmPC) of lenvatinib, sorafenib and cabozantinib all have warnings about use in patients with prolonged QT intervals under the SmPC section of 'Special warnings and precautions for use.'"	Ensure reader understands that caution in using cabozantinib in a patient with a prolonged QT interval would also apply to lenvatinib and sorafenib and if a clinician would not use cabozantinib in a patient with a prolonged QT interval then they should not be using lenvatinib or sorafenib either.	The EAG has amended the text as requested.

conclude lenvatinib and sorafenib could be used in patients with a prolonged QT interval but cabozantinib cannot.			

Issue 4 Incorrect table

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 46, Table 12: These figures used for CCO1 are incorrect.	The company submission table (used in the EAG report) contained the incorrect figures for the PFS per BIRC ITT population (table 11, CS). This was noticed by the company during the clarification questions and the correct table to include in the EAG report can be found in Response C2 (Table 12) of the clarification questions.	Incorrect figures.	The EAG agrees. The correct figures from the clarification response have now been included in Table 12, including the removal of AIC designations, which were present in the CS but absent from the clarification response, and some CCO2 data that were also different between these tables. Please ensure that you are happy with the amended AIC marking.

Issue 5 Incorrect table

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 49, Table 14: These figures used for CCO1 are incorrect.	The company submission table (used in the EAG report) contained the incorrect figures for the OS per BIRC ITT population (table 11, CS). This was noticed by the company during the clarification questions and the correct table to include in the EAG report can be found in Response C2 (Table 13) of the clarification questions.	Incorrect figures.	The EAG agrees. The correct figures from the clarification response have now been included in Table 14, including the removal of AIC designations, which were present in the CS but absent from the clarification response. Please ensure that you are happy with the amended AIC marking.

Issue 6 Incorrect data from trial put into table

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 56, Table 18: The row which describes the number of patients by geographical region for participants in the NCT01811212 trial is aligned to Europe but it should be aligned to North America (USA and Canada) as all participants were from the USA.	Correct the misalignment of the 25 (100%) of patients in Table 18 of the EAG report so that it aligns to North America (USA and Canada) instead of Europe from the NCT01811212 study.	Ensure reader does not misinterpret the data.	The EAG agrees. The numbers have been re-aligned as requested.

Issue 7 Incorrect spelling

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 62: There is a spelling error in this sentence "Section 5.1 describes and critiques the company's SLR of existing ec onomic analyses of treatments for RAI-refractory DTC."	"Section 5.1 describes and critiques the company's SLR of existing **economic** analyses of treatments for RAI-refractory DTC."	Minor adjustment for clarification	The EAG disagrees – this typo does not feature in the EAG report submitted to NICE. The EAG report has not been amended.

Issue 8 Incorrect description of SLR results

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 63: There is an incorrect statement that there were five CUA studies and one CEA. There were 3 CUA and 3 CEA. "Five of the studies were cost-utility analyses which reported outcomes in terms of the incremental cost per quality-adjusted life year (QALY) gained; the sixth study, Carlson *et al.*,[45] does not report costs or incremental cost-effectiveness ratios (ICERs) and therefore should have been excluded according to the eligibility criteria for the review (see CS Appendix G,[25] Table 37)."	"**Three of the six studies were cost-utility analyses, with the remaining three being cost-effectiveness analyses. Five of the studies** reported outcomes in terms of the incremental cost per quality-adjusted life year (QALY) gained; the sixth study, Carlson *et al.*,[45] does not report costs or incremental cost-effectiveness ratios (ICERs) and therefore should have been excluded according to the eligibility criteria for the review (see CS Appendix G,[25] Table 37)."	Minor adjustment for clarification	The EAG disagrees. The study by Carlson is not an economic evaluation as it does not report costs – it is therefore incorrect to label it as a cost-utility analysis and this study should have been excluded from the company's review of economic evaluations. All of the other five identified studies report economic results in terms of the incremental cost per QALY gained – therefore we have referred to these as cost-utility analyses. The text in the EAG report is already factually accurate and therefore has not been amended.

Issue 9 Incorrect TTD data description

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 68: Time on treatment was modelled on TTD data linked to date of last observed dose. "Time on treatment is modelled using an exponential distribution fitted to data on TTD, linked to the date of progression"	"Time on treatment is modelled using an exponential distribution fitted to data on TTD, linked to the **date of last observed dose**. **The Kaplan-Meier data was incorrect and was based on the TTD linked to the date of progression, however this was updated in the clarification response to question A15."**	Following consultation with the company's global team, the original TTD coefficients were based on the last observed dose. However, the Kaplan Meier curves were incorrect as these were based on TTD date of progression and were subsequently updated in the clarification response to question A15.	The EAG agrees. The text has been amended to read: *"Time on treatment is modelled using an exponential distribution fitted to data on TTD, linked to the date of the last known dose."*

Issue 10 Incorrect reference

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 71: The values in the following statement are derived from the text and table in response B14, not B9.	"The company's sensitivity analysis assessing the common treatment effect assumption shows that the treatment effect varies from 0.65 to [redacted] when changing the assumption	Minor adjustment for clarification	The EAG agrees. The text has been amended to refer to question B14.

"The company's sensitivity analysis assessing the common treatment effect assumption shows that the treatment effect varies from 0.65 to [redacted] when changing the assumption from the common treatment effect to no treatment effect in crossover patients (see clarification response,[17] question B9)."	from the common treatment effect to no treatment effect in crossover patients (see clarification response,[17] question **B14**)."		

Issue 11 Factual inaccuracy

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 78: The model and KM curve reflect the CCO2 data-cut for cabozantinib so Figure 22 is correct. CCO1 is only selected for OS for the BSC arm.	~~"The plot shown in Figure 22 of the company's clarification response is incorrect as the Kaplan-Meier estimates reflect the CCO1 data-cut whilst the models reflect the CCO2 data-cut."~~	Removal of sentence for accuracy	The EAG agrees with the company that there is an error in the EAG's footnote. However, there is also an error in Figure 22 of the company's clarification response which needs to be

"The plot shown in Figure 22 of the company's clarification response is incorrect as the Kaplan-Meier estimates reflect the CCO1 data-cut whilst the models reflect the CCO2 data-cut."			explained. This explanation is given in the company's response to follow-up questions from the EAG (received 16/11/22). The EAG's footnote has been amended to read: "*The Kaplan-Meier plot of TTD presented in Figure 27 of the CS does not reflect the TTD survival models as it includes the PFS cap. The plot shown in Figure 22 of the company's clarification response is incorrect as the Kaplan-Meier estimates reflect TTD linked to progression rather than TTD linked to the observed last dose."*

Issue 12 Spelling error

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 85: No space between COSMIC-311 and trial. "Costs related to the management of AEs were	"Costs related to the management of AEs were based on the frequency of individual Grade 3/4 TEAEs with an incidence ≥5% observed in either the cabozantinib arm or	Minor amendment	The EAG agrees. This typo has been fixed.

based on the frequency of individual Grade 3/4 TEAEs with an incidence ≥5% observed in either the cabozantinib arm or placebo arm of the ITT population of the COSMIC-311trial at CCO1”	placebo arm of the ITT population of the **COSMIC-311 trial** at CCO1”		

Issue 13 Factual inaccuracy

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**EAG response**
Page 86: Health state utilities were included in the original model PSA, however there was an unknown #REF error in place of the SE. “These parameters are uncertain and should have been included in the PSA.”	-“These parameters **were intended for inclusion in the original models PSA, however a #REF error was found in place of the SE, subsequently stopping utilities from being included.”**	Amendment for clarification. The company had full intention of including utilities for PD and PFS health states in the PSA.	This is not a factual inaccuracy, but for the sake of clarity, the following additional text has been added to the EAG report. *“At the factual accuracy check stage of the appraisal, the company clarified that they had intended to include the utility values in the PSA..”*

Issue 14 Factual inaccuracy

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 94: The company amended issue (iii) in the revised model. General population mortality was amended to match ONS National life tables. "The company's revised model provided as part of their clarification response includes amendments to address all of these issues, except for issue (iii)." Page 109: General population mortality was amended in the revised model to match ONS National life tables "The model errors described in critical appraisal point [1]	"The company's revised model provided as part of their clarification response includes amendments to address all of these issues, **however the EAG preferred a different approach**" "The model errors described in critical appraisal point [1] of Section 5.3.5 have been resolved, **however the EAG opted for an alternative approach to issue (iii).**"	The EAG's model updates general population mortality in an alternative approach to the company, however, this issue was addressed in the revised model with ONS National life tables.	This is not a factual inaccuracy. The life table inputs were amended in the revised model, but the approach used to estimate general population mortality risk was not amended. Both the company's original and revised models apply the same calculations which assume that: (a) mortality risks differ between men and women and (b) that the same proportion of surviving patients are men and women. As both assumptions cannot be simultaneously true, the EAG considers this to be an error.

of Section 5.3.5 have been resolved, except for issue (iii)."			The EAG report has not been amended.

Issue 15 Factual inaccuracy

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 101: In B6, the company states the revised model uses CCO1 OS for BSC arm, removing the 5-year constraint, and a smooth decline has replaced the sharp drop. "The EAG asked the company whether their experts considered the exponential model including the 5-year death assumption to be plausible (see clarification response,17 question B6); however, the company did not provide a response on this issue. Despite this, the	"The EAG asked the company whether their experts considered the exponential model including the 5-year death assumption to be plausible (see clarification response,[17] question B6). **The company responded by incorporating CCO1 OS for the BSC arm, removing the need for the 5-year OS constraint. This amendment removed the clinically implausible vertical drop and replaced it with a smooth decline, which the EAG notes is more clinically realistic.**"	Text amendment to clarify the Company's response to clarification question B6	The text is not factually inaccurate. The company's proposed text in bold on the left explains the company's rationale for removing the 5-year death assumption. However, the EAG's original text is factually accurate - the company's clarification response did not provide a response about whether the company's clinical experts believed that the OS curve based on CCO2 with the 5-year death assumption was clinically plausible. The approach used in the

EAG notes that this structural assumption results in a vertical drop in the survivor function for BSC which is not clinically realistic.”			company’s revised model is explained elsewhere in the report. The EAG report has not been amended.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **Wednesday 25th January 2023**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we

received, and are not endorsed by NICE, its officers or advisory committees.

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Ipsen Ltd
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Uncertainty around the effect of cabozantinib and best supportive care (BSC) on	Yes	The EAG consulted two clinical experts to provide their estimates for the expected survival of cabozantinib and BSC patients at 2, 5 and 10 years. This is shown below and taken from Table 37 of the EAG report.

overall survival		[redacted] The estimates for survival for BSC at 5 years in the above table from the company's clinical advisors comes from an advisory board conducted with three clinical experts in August 2022.[1] In addition to the advisory board, in preparation for the NICE submission, the company had also asked four clinical experts in individual meetings via a [redacted] what they would expect survival to be for cabozantinib and BSC at 2, 5 and 10 years as shown in Table 1 below. These meetings occurred between 14/04/22 and 07/07/22. One of the experts who was interviewed by the [redacted] was also interviewed by the [redacted] in November 2022 to inform the Scottish Medicines Consortium (SMC) submission currently in development, as such the mean average of the estimates provided by this expert who was interviewed twice is shown in Table 1. To strengthen the clinical expert evidence base, rather than use the company advisory board input alone, the company believe it is appropriate to pool all available evidence to derive a more robust elicitation and consensus. On this basis, Table 1 incorporates the EAG expert estimates, input from the company advisory board and company advisors via the individual meetings to provide an overall mean average estimate of seven different clinical experts. Note: two of the individuals from the company individual meetings also participated in the advisory board held by the company. The

		company believes these overall mean average clinical expert estimates provide more precision and reduce the uncertainty of survival expected at 2, 5 and 10 years and therefore should be used to inform decision making by the NICE committee. The overall mean average clinical expert estimates for cabozantinib at 2, 5 and 10 years of [redacted] and [redacted] align with the company modelled cabozantinib survival of [redacted] and [redacted] this approach is maintained and agreed in the EAG preferred analysis. For BSC, the mean clinical experts estimates at 2, 5 and 10 years of [redacted] and [redacted] align more closely with the company's modelling approach (including the 5-year BSC death assumption) than the EAG preferred analysis (excluding the 5-year BSC death assumption). The company models 0% survival from 5 years rather than an overestimate of [redacted] and [redacted] modelled in the EAG analysis at 5 and 10 years, respectively. Therefore, the company maintain that the inclusion of the 5-year BSC death assumption is appropriate compared to excluding this assumption. Furthermore, the survival estimates from the experts align even closer to the CCO1 data cut for BSC which was provided by the company in response to B1 and B6 of the clarification response. However to alleviate the EAGs concern with application of a 5-year death assumption, the company has performed an analysis on second-line only patients from COSMIC-311. In England and Wales there are no second-line options after initial TKI therapy in radioiodine refractory (RAIR) DTC patients. Therefore, cabozantinib will be prescribed as a second-line therapy. Approximately 75% of patients in the COSMIC-311 trial had only one prior TKI. The company conducted an analysis of pure second-line patients from the COSMIC-311 trial to model effectiveness in this specific population and subsequently this analysis has been used to inform the base case analysis. In addition, a blended survival scenario analysis incorporating clinical expert opinion into CCO2 data has been conducted. The results of all these analyses are summarised in Table 1 below and a more detailed description is provided below the table of the analyses for the pure second-line population and the blended survival analysis for the whole COSMIC-311 trial population. **Table 1: Clinical experts estimates for OS at 2, 5 and 10 years for cabozantinib and BSC**

Advisor estimates	Cabozantinib			BSC		
	2 years	**5 years**	**10 years**	**2 years**	**5 years**	**10 years**
EAG Advisor 1	63%	35%	0%	50%	10%	0%
EAG Advisor 2	45%	25%	0%	35%	13%	0%
Company's clinical advisors[1]	-	-	-	-	[redacted]	-
Company Advisor 1*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company Advisor 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company Advisor 3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company Advisor 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean average across EAG and company advisors	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company's model with 2L base case analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company's model (including BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

		5-year death assumption)						
		Company's model (excluding BSC 5-year death assumption)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Company's model (using CCO2 for cabozantinib and CCO1 for BSC)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Company's model with blended survival scenario analysis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Note: Mean average calculated where ranges provided by advisors.

**Mean average calculated from two interviews conducted with same clinical advisor.*

***Note: The proportion of patients modelled to be alive at 2 years on BSC in the company model excluding BSC 5-year death assumption is [redacted] rather than [redacted] reported in the EAG report.*

Abbreviations: BSC – best supportive care; CCO1 – Clinical cut-off 1; CCO2 – Clinical cut-off 2; *EAG – external assessment group; OS – overall survival*

Pure second-line analysis

For the analysis of the second-line only data from COSMIC-311, the company compared the PFS curves based on comparison to the curves selected by three UK clinicians[1] in a previous advisory board held in August 2022, the goodness of fit (AIC and BIC), and visual inspection against the observed KM data. Table 2 presents a summary of the second line analyses for PFS, OS and TTD. The hazard ratio (HR) stratified was [redacted]), The median duration of PFS were [redacted] months in the cabozantinib arm and [redacted] months for the placebo arm. The KM estimates for median duration of OS were not estimable in either arm. The median TTD were [redacted] months and [redacted] months for cabozantinib and placebo respectively.

Table 2: PFS, OS and TTD CCO2 pure second line

	CCO2 pure 2L PFS (N = 191)			**CCO2 pure 2L OS (N = 191)**				**CCO2 pure 2L TTD (N = 191)**	
	Cabozantinib (N = 126)	**Placebo (N = 65)**		**Cabozantinib (N = 126)**	**Placebo (N = 65)**	**Placebo RPFST adjusted (N = 65)**		**Cabozantinib (N = 126)**	**Placebo (N = 65)**
Number (% of subjects)				**Number (% of subjects)**			**Number (% of subjects)**		
Event	[redacted]	[redacted]	Event	[redacted]	[redacted]	[redacted]	Event	[redacted]	[redacted]
Death	[redacted]	[redacted]	-	-	-	-	-	-	-

		Progressive disease	[redacted]	[redacted]	-	-	-	-	-	-	-
		Duration of PFS (months)				**Duration of OS (months)**			**Duration of TTD (months)**		
		Median (96% CI)	[redacted]	[redacted]	Median (96% CI)	[redacted]	[redacted]	[redacted]	Median (96% CI)	[redacted]	[redacted]
		Hazard ratio (96% CI; stratified)	[redacted]		**Hazard ratio (96% CI; stratified)**	[redacted]		[redacted]	**Hazard ratio (96% CI; stratified)**	[redacted]	
		Hazard ratio (96% CI; unstratified)	[redacted]		**Hazard ratio (96% CI; unstratified)**	[redacted]		[redacted]	**Hazard ratio (96% CI; unstratified)**	[redacted]	

Abbreviations: CCO2 – Clinical cut-off 2; CI – confidence interval; NA – not applicable; OS – overall survival; PFS – progression free survival; TTD – time to discontinuation

Figure 1 and Figure 2 show PFS data fitted and extrapolated using the standard parametric models for the cabozantinib and BSC arms, respectively.

From the previous advisory board, three UK clinicians selected PFS distributions that had 0% PFS at 5 years (Table 3). In the second-line only analysis, only the Weibull and Gompertz distributions have 0% at 5 years for cabozantinib and BSC. Based on the goodness of fit statistics and visual inspection against the observed KM data, the Weibull distribution has the lowest AIC and BIC and best fit to KM data compared with Gompertz (Table 4). Therefore, for the base case, the Weibull distribution was selected. Upon applying PFS in the model a rule was also applied whereby the PFS curve could not exceed the OS curve for each treatment.

Table 3: Percentage of patients progression free for second-line only

	3 months		6 months		1 year		2 years		5 years	
	Cabo	BSC	Cabo	BSC	Cabo	BSC	Cabo	BSC	Cabo	BSC
Exponential	[redacted]									
Weibull										
Gompertz										
Log-logistic										
Log-normal										
Generalized gamma										

Abbreviations: BSC – best supportive care

Table 4: PFS AIC and BIC for second-line only

Distribution	Cabozantinib			BSC			Total Sum
	AIC	BIC	Sum	AIC	BIC	Sum	
Exponential	[redacted]						
Weibull							
Gompertz							

Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Generalized gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC – Akaike information criterion; BIC – Bayesian information criterion; BSC – best supportive care; PFS – progression free survival

Figure 1: Cabozantinib PFS curves for second-line only

[redacted]

Abbreviations: KM – Kaplan Meier; PFS – progression free survival

Figure 2: BSC PFS curves for second-line only

Abbreviations: BSC – best supportive care; KM – Kaplan Meier; PFS – progression free survival

For OS, the company compared the OS curves based on comparison to the curves selected by three UK clinicians[1] in a previous advisory board held in August 2022, the goodness of fit (AIC and BIC), and visual inspection against the observed KM data.

All OS curves overestimate the survival expectations the clinical experts had in the advisory board (0% at 5 and 10 years). All OS curves overestimate the mean survival expectations of all clinical experts (advisory board and EAG clinical experts (5% at 5 year and 0% at 10 years) (Table 1). Therefore, the company selected the exponential curve based on goodness of fit (AIC and BIC) and visual inspection (Table 5, Figure 3 and Figure 4)

Table 5: OS AIC and BIC for second-line only

Distribution	Cabozantinib			BSC			Total Sum
	AIC	BIC	Sum	AIC	BIC	Sum	
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Generalized gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC – Akaike information criterion; BIC – Bayesian information criterion; BSC – best supportive care; OS – overall survival

Figure 3: Cabozantinib OS curves for second-line only

Abbreviations: KM – Kaplan Meier; OS – overall survival

Figure 4: BSC OS curves for second-line only

		 *Abbreviations: BSC – best supportive care; KM – Kaplan Meier; OS – overall survival* For TTD, the company compared TTD curves based on closeness to PFS curve selected (Weibull), goodness of fit (AIC and BIC; Table 6) and visual inspection against the observed KM data (Figure 5). The two curves that were closest to the PFS curve were Gompertz and generalised gamma. This aligns with the SmPC for cabozantinib wording that "*patients should continue treatment until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs*".[2] The TTD is based on the date of the last dose. Based on the goodness

of fit and visual inspection against observed KM data, the generalised gamma distribution has the lowest AIC and BIC and best fit to KM data compared with Gompertz (Table 6).

Table 6: TTD AIC and BIC for second-line only

Distribution	Cabozantinib		
	AIC	BIC	Sum
Exponential	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]
log-logistic	[redacted]	[redacted]	[redacted]
log-normal	[redacted]	[redacted]	[redacted]
Generalized gamma	[redacted]	[redacted]	[redacted]

Abbreviations: AIC – Akaike information criterion; BIC – Bayesian information criterion; TTD – time to discontinuation

Figure 5: Cabozantinib TTD curves for second-line only

Abbreviations: KM – Kaplan Meier; TTD – time to discontinuation

The results and scenario results for the second line only analysis can be found in the 'Summary of changes to the company's cost-effectiveness estimate(s)' section below.

Blended survival analysis

Ipsen conducted a blended survival scenario analysis incorporating clinical expert opinion into CCO2 data. In a recent paper[3] a new approach was introduced incorporating expert elicitation to form a blended survival curve. The blended survival approach involves creating a new survival function where the observed data gradually approaches the parametric model informed by external data over the extrapolation period. Since the estimates from the survival extrapolation of the OS data for the BSC arm of COSMIC-311 excluding the BSC 5-year death assumption exceeds the

		mean average estimates from the panel of experts (two EAG advisors and five consulted by Ipsen) as indicated in Table 1. The blended survival approach was used as a scenario to seek closer alignment between the survival and the averaged estimates of the experts. Blended survival curves were not used for the cabozantinib arm as the expert estimates of survival probabilities are similar to what the health economic model estimates. To create the blended survival curve, the example code provided in the Che et al. 2022[3] paper was adapted. To align with the original approach for treatment crossover, OS for the BSC arm adjusted according to the Rank-Preserving Structural Failure Time (RPFST) method was used. As mentioned in the paper, the best fit possible to the observed information was suggested, hence a piecewise exponential model (with time intervals of 2 months) was fitted for the observed data (i.e. COSMIC-311 BSC arm – RPFST adjusted OS). For the extrapolation using expert opinion, the mean average estimates for probability of survival for the BSC arm at given times were used as presented in Table 1. The example code from the paper only allowed for two time points, including the timepoint where survival probability equates to zero. To better fit the data estimates provided by the experts, an additional timepoint was added to the code, as such all three timepoints from Table 1 (2, 5 and 10 years) are included. A number of different distributions were tested. The Weibull distribution provided a good fit among the models tested (standard parametric) and a conservative survival estimate over the long term (Figure 6). As such the Weibull was chosen to be the base case scenario for the external data extrapolation. However, the Weibull still slightly overestimated survival probabilities at the 3 given timepoints (Table 1). To explore the uncertainty around the estimation, the base case scenario utilised N=70 patients for scenario analyses the number of patients was changed by ±20%. **Figure 6. All models fitting for expert opinion survival curve**

		 **Figure 7. Weibull fit for expert judgement survival curve**

		 The blended survival curve is obtained by blending the survival curve extrapolated from observed data and survival curve informed by expert opinion between timepoints using a weighting function. There are two sets of parameters for the weighting function, including the blending interval, and beta random variable with parameters alpha and beta. For the blending interval, the base case scenario utilised 14 months, as observed data is only available up until 13.6 months and the initial data point for the expert estimates was 24 months. A scenario

analysis using 14 months and blending to the second timepoint for the expert estimates (i.e. 60 months was used). In the base case scenario, the beta random variable with parameters alpha=3 and beta=3 was used, this gives a steady blending weight. A scenario with a faster blending rate (alpha=2 and beta=5) was also tested. Table 7 presents the base case and scenario values.

Table 7. Assumptions for the base case and scenarios

	Base case	Scenario
Model for observed data	Piecewise exponential model	NA
Model for expert opinion survival curve	Weibull	NA
Uncertainty of external data	N=70	N=70 ±20%
Blending interval	14 months to 24 months	14 months to 60 months
Parameter for weight function	Alpha = 3 Beta = 3	Alpha = 2 Beta = 5

Abbreviations: NA – not applicable

The base case scenario for the blended survival analysis produced mean survival estimates of [redacted] and [redacted] at 2, 5, and 10 years respectively. The graphical representation of the blended curve is displayed in Figure 8.

Figure 8: Blended RPSFT-adjusted BSC curve (Base case scenario)

Abbreviations: BSC – best supportive care; KM – Kaplan Meier; RPSFT – Rank-Preserving Structural Failure Time

As detailed in Table 7, scenarios including changing the blending interval, changing the blending rate, using lower uncertainty expert estimates and using higher uncertainty expert estimates were also performed. The graphical representations of these curves and the mean survival estimates are presented in Table 8 and Figure 9 to Figure 12. The cost-effectiveness results for the blended scenario analysis can be found in the Scenario analysis section below.

Table 8: Assumptions for the base case and scenarios

Scenario	2 years	5 years	10 years
Blending interval: 60 months	[redacted]	[redacted]	[redacted]
Blending rate: Alpha = 2, Beta = 5	[redacted]	[redacted]	[redacted]
Low uncertainty	[redacted]	[redacted]	[redacted]
High uncertainty	[redacted]	[redacted]	[redacted]

		Figure 9: Blended RPSFT-adjusted BSC curve (Blending interval: 60m)  *Abbreviations: BSC – best supportive care; KM – Kaplan Meier; RPSFT – Rank-Preserving Structural Failure Time* **Figure 10: Blended RPSFT-adjusted BSC curve (Blending rate: Alpha=2, Beta=5)**  *Abbreviations: BSC – best supportive care; KM – Kaplan Meier; RPSFT – Rank-Preserving Structural Failure Time*

		Figure 11: Blended RPSFT-adjusted BSC curve (Low uncertainty)  *Abbreviations: BSC – best supportive care; KM – Kaplan Meier; RPSFT – Rank-Preserving Structural Failure Time* **Figure 12: Blended RPSFT-adjusted BSC curve (High uncertainty)**  *Abbreviations: BSC – best supportive care; KM – Kaplan Meier; RPSFT – Rank-Preserving Structural Failure Time*
Uncertainty	No	The company maintain that the Fordham et al.[4] utilities are the most appropriate utility source for the base case analysis due to the lack of validity in the COSMIC-311 health related quality-of-life (HRQoL) data as raised in B20 in

around the most appropriate health state utility values		the clarification response document, and agreed with by the EAG in their report. Please note: to complete the response to B18 in the clarification response, the latest data-cut CCO2 was used for the utility analysis from COSMIC-311 such that no further data-cuts are planned. The utility values used in the company model are adjusted regression values (PFS:0.87, PD:0.52). The EAG have noted that the general population utility value of 0.82 is lower than that of the PFS value of 0.87. To correct for this, an age-adjusted general population utility cap was applied in the revised company model as per B17 in the clarification response document which was also implemented in the EAG adapted model (the model which has been used by the company from herein) so that the utility value for PFS can never exceed general population utility. This is in line with the opinion of one of the EAG's advisor who stated that "for patients who are fit, progression-free and are not experiencing toxicity, they would expect HRQoL to be similar to general population levels", with a second advisor adding "whilst there may be negative psychological impacts on a patient's HRQoL due to their diagnosis of differentiated thyroid carcinoma (DTC), in terms of physical impacts, HRQoL would be similar to that of the general population", as per the EAG report. The second critique by the EAG is that the company model uses adjusted regression values instead of the mean observed values in the Fordham et al. study[4]. The utility values used in the company model are adjusted for educational qualification level and EQ-5D-3L (usual activities and anxiety/depression) ratings using UK norms regression values (PFS:0.87, PD:0.52), instead of the observed mean values (PFS:0.80, PD:0.50). The study describes how the sample size in the valuation may have affected the representativeness of the results for the general population. The EQ-5D-3L data suggested that the sample were relatively healthier than the general population, in addition to the sample overrepresenting higher education completers and married people status when compared to UK census data. Fordham et al. performed a regression analysis, adjusting for these imbalances, to produce utility values which are closer to those expected from a more normative UK sample. The company propose to continue the use of these adjusted utility values which significantly predicted utility set to UK normative values. Furthermore, the decrement of 0.35 between the adjusted utility values for PFS (0.87) and PD (0.52) were considered expected and plausible by UK clinicians[1] and the EAG's clinical advisors (as stated in the EAG report), respectively.

		Considering all the above, the company maintain that the adjusted utility values from Fordham et al. are the most appropriate to use in the base case analysis with the age-adjusted general population utility cap.
Issues relating to resource use and costs	No	The company agree with the EAG's conclusion that costs of post-progression cabozantinib should be included in the economic analysis to reflect the intention for cabozantinib to be used in line with its licence, which permits continued treatment after progression for patients who are still clinically benefitting from treatment. The company agree with the approach taken in the EAG adapted model, removing the stopping rule which caps time to treatment discontinuation (TTD) by PFS. In addition, the company agrees with the use of the Weibull distribution for the previous base case analysis to model TTD. The company has followed the same decision process for selecting the TTD curve in the new base case with pure second line data.
	No	The company maintains that using relative dose intensity (RDI) instead of compliance is the most appropriate way of deriving the true cost per cycle of cabozantinib due to uncertainty around the validity of the compliance figure and methods in past NICE appraisals. The analysis of compliance used to inform the economic model is not stated in the clinical study report or any of its addendums and was calculated based on CCO1 patient level data, however RDI was analysed and included in the clinical study report for CCO2. The compliance figure was based on data for CCO1 only and was not rerun for CCO2. CCO1 has very limited follow up of only 6.24 months (median). Whereas RDI is based on CCO2 data. Data with longer follow up should be used where possible for all inputs in the economic model including the dose intensities of the treatments. In addition, previous NICE appraisals have been consistent in the inclusion of RDI in the economic model regardless of whether the medicine was linear pricing per mg or flat price per mg. The most relevant example is in TA535 for lenvatinib and sorafenib in first line radioactive iodine (RAI) refractory DTC. Lenvatinib is flat priced with both the 4 mg and 10 mg priced the same. No issues were raised in this appraisal on the use of RDI to calculate the "true cost" of lenvatinib[5]. Therefore, it would be unreasonable and inconsistent to apply a different method in this appraisal. In past appraisals of cabozantinib, RDI has also been used to adjust for the true cost per cycle of treatment such as the recent TA849, cabozantinib for previously treated advanced hepatocellular carcinoma[6]. The company believes a

		consistent approach should be adopted in applying the RDI, as done previously and accepted, to treatments in this appraisal. Therefore, the company believe that consistency should be maintained with NICE/EAG standards adhered to such that the most up to date data should be used where possible/relevant and between past appraisals. On this basis, application of RDI is most appropriate in the base case. Furthermore, the EAG preferred assumption of implementing wastage once per patient is supported by the company and has been kept in the base case analysis.
	No	The company deem the EAG scenario analysis assuming double the frequency of ECGs required for patients receiving cabozantinib i.e. from every three months rather than six months, inappropriate and insufficiently substantiated. The Summary of Product Characteristics (SmPC) for cabozantinib states *"when using cabozantinib, periodic monitoring with on-treatment ECGs should be considered."*[7] This wording is also included in the SmPCs for lenvatinib and sorafenib which belong to the same class of medicine as cabozantinib (tyrosine kinase inhibitors - TKIs). One of the clinical experts consulted by the EAG is from the Clatterbridge Cancer Centre NHS Foundation Trust. This hospital's protocol for monitoring lenvatinib for DTC states an ECG should be done prior to treatment and then as clinically indicated[8]. Due to ambiguity of wording in SmPC and EAG clarification, we believe it's best to follow the clinical opinion from experts at the company's previous advisory board. Feedback from experts at the company's advisory board concluded ECGs would on average be performed every six months. Therefore the company believes, that on balance taking into account expert opinion and an NHS Trust guideline of one of the EAG clinical experts being ambiguous, the frequency of ECG monitoring for cabozantinib used in the health economic model should follow the advice of clinical opinion in the company's advisory board. Therefore, ECGs every 3 months is not appropriate.
	No	The company deem the EAG scenario analysis assuming that patients receiving BSC do not to undergo any CT scans inappropriate. Ipsen notes that in nearly all NICE technology appraisals for thyroid cancer CT scans have been applied to patients receiving BSC.

		In TA535[5] and TA742[9] the EAG assumed or accepted that CT scans were performed every 3 months for both the intervention and BSC. In TA516 the EAG assumed the frequency of CT scans was once every 3 months for the intervention and once every 6 months for BSC[10]. Therefore the company believes for consistency with other appraisals it should be assumed the CT scans are performed with equal frequency in the intervention and BSC arms in the base case analysis, as per the company and EAG preferred analysis, however a scenario considering no CT scans for BSC is not appropriate.
	Yes	The company has considered the inclusion of concomitant medications in the economic analysis. Concomitant medications are defined in the COSMIC-311 trial as those that stop or continue on or after date of first dose through the end of the safety observation period. No data are available to compare the proportions of the concomitant medicines patients were taking before entry into the trial and how this changed during the trial. As shown in Table 9, the receipt of concomitant medications between the treatment arms, cabozantinib and placebo, were balanced. Note: only data for CCO1 data cut are available. **Table 9: Concomitant medications received by treatment arm based on CCO1**

	CCO1	
Parameter	**Cabozantinib (N=125) n (%)**	**Placebo (N=62) n (%)**
Number of subjects that received at least one concomitant medication	[redacted]	[redacted]

Abbreviations: CCO1 – Clinical cut-off 1

		Concomitant medications that had a ≥ 10% higher incidence in the cabozantinib only arm compared with the placebo arm by decreasing frequency of between-arm difference were loperamide, amlodipine, urea, calcium, paracetamol (acetaminophen), and clobetasol. There is no information on the duration of use of these medicines and all are generics with a very low cost. Therefore, inclusion in the economic analysis would have a minimal impact on the incremental cost-effectiveness ratio (ICER) and therefore the company maintain not including these costs in the base case analysis.

References

(1) Ipsen. *IPSEN HTA Advisory Board for Differentiated Thyroid Cancer. [Data on File]*; 2022.
(2) Cabometyx-Epar-Product-Information_en.Pdf. https://www.ema.europa.eu/en/documents/product-information/cabometyx-epar-product-information_en.pdf (accessed 2022-08-19).
(3) Che, Z.; Green, N.; Baio, G. *Blended Survival Curves: A New Approach to Extrapolation for Time-to-Event Outcomes from Clinical Trials in Health Technology Assessment*. https://doi.org/10.1177/0272989X221134545.
(4) Fordham, B. A.; Kerr, C.; de Freitas, H. M.; Lloyd, A. J.; Johnston, K.; Pelletier, C. L.; Tremblay, G.; Forsythe, A.; McIver, B.; Cohen, E. E. Health State Utility Valuation in Radioactive Iodine-Refractory Differentiated Thyroid Cancer. *Patient Prefer Adherence* **2015**, *9*, 1561–1572. https://doi.org/10.2147/PPA.S90425.
(5) NICE. Lenvatinib and Sorafenib for Treating Differentiated Thyroid Cancer after Radioactive Iodine [ID1059], 2018. https://www.nice.org.uk/guidance/ta535/documents/committee-papers.
(6) NICE. Cabozantinib for Previously Treated Advanced Hepatocellular Carcinoma [TA849] [Internet]. 2022. Available from: Https://Www.Nice.Org.Uk/Guidance/Ta849.
(7) EMA. Summary of Product Characteristics, 2022. https://www.ema.europa.eu/en/documents/product-information/cabometyx-epar-product-information_en.pdf (accessed 2022-08-10).
(8) The Clatterbridge Cancer Centre NHS Foundation Trust. Systemic Anti-Cancer Treatment Protocol – Lenvatinib (Lenvima®) Differentiated Thyroid Cancer Protocol. PROTOCOL REF: MPHALENVHN (Version No: 1.2). Available at: Https://Www.Clatterbridgecc.Nhs.Uk/Application/Files/8816/1349/6317/Lenvatinib_LENVIMA_Differentiated_Thyroid_Cancer_Protocol_V1.2.Pdf.
(9) NICE. Selpercatinib for Treating Advanced Thyroid Cancer with RET Alterations [ID3744], 2021. https://www.nice.org.uk/guidance/ta742/documents/committee-papers.

(10) NICE. Cabozantinib and Vandetanib for Treating Unresectable Locally Advanced or Metastatic Medullary Thyroid Cancer [ID56], 2017. https://www.nice.org.uk/guidance/ta516/documents/committee-papers.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
N/A	N/A	N/A	N/A

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Issue 1 – Uncertainty around the effect of cabozantinib and BSC on overall survival	Previous ICER including QALY weighting (£/QALY): £20,289	Analysis of the pure second-line population (Settings G37) from COSMIC-311 was incorporated into the health economic model. • PFS distribution: Weibull (Updated EAG code such Weibull distribution was no longer hardcoded in Trace (Cabozantinib I9:I608) • OS distribution: Exponential • TTD distribution: generalised gamma • RDI included (Settings G33)	ICER including QALY weighting (£/QALY): £19,208 Change from original base-case ICER: £1,081

		• Utility values from Fordham et al. adjusted (Quality of life inputs D12:D13)	
Issue 1 – Uncertainty around the effect of cabozantinib and BSC on overall survival	N/A	Added in Blended Survival analysis into model (Settings G39:G40). This involved an edit to EAG formulae in EAG_GenPopAndTrace V7:V427 to ensure the selected survival estimates were pulled through into the EAG corrected trace.	N/A
Company's base case following technical engagement (or revised base case)	[redacted]	[redacted]	ICER including QALY weighting (£/QALY): £19,208

Sensitivity analyses around revised base case

Scenario analysis

Table 10 details deterministic scenario analysis results for cabozantinib versus BSC.

Table 10: Deterministic scenario analysis results (pure second-line population)

Description	Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs excluding QALY weighting	ICER including QALY weighting (£/QALY)	ICER (£) excluding QALY weighting (£/QALY)
Base case *(3.5%, exponential OS, Weibull PFS, gen-gamma TTD, RDI, adjusted utility values)*	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,208	23,050
Discount rate: 0%	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,047	21,656
Discount rate: 5%	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,687	23,624
Compliance	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,508	25,810
5-year OS constraint	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,821	21,386
OS curve: lognormal	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	15,376	18,451
OS curve: Log-logistic	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	25,182	30,218

Description	Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs excluding QALY weighting	ICER including QALY weighting (£/QALY)	ICER (£) excluding QALY weighting (£/QALY)
PFS curve: Lognormal	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	15,787	18,944
PFS curve: Generalised gamma	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,251	21,901
TTD curve: Exponential	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,241	26,690
TTD curve: Gompertz	BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	-
	Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,037	22,844

Abbreviations: BSC – best supportive care; ICER – incremental cost-effectiveness ratio; LYG – life years gained; OS – overall survival; PFS – progression free survival; QALY – quality adjusted life years; RDI – relative dose intensity; TTD – time to treatment discontinuation

Probabilistic results.

Base case

Table 11: Probabilistic scenario analysis results – Base case (pure second-line population)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs excluding QALY weighting	ICER including QALY weighting (£/QALY)

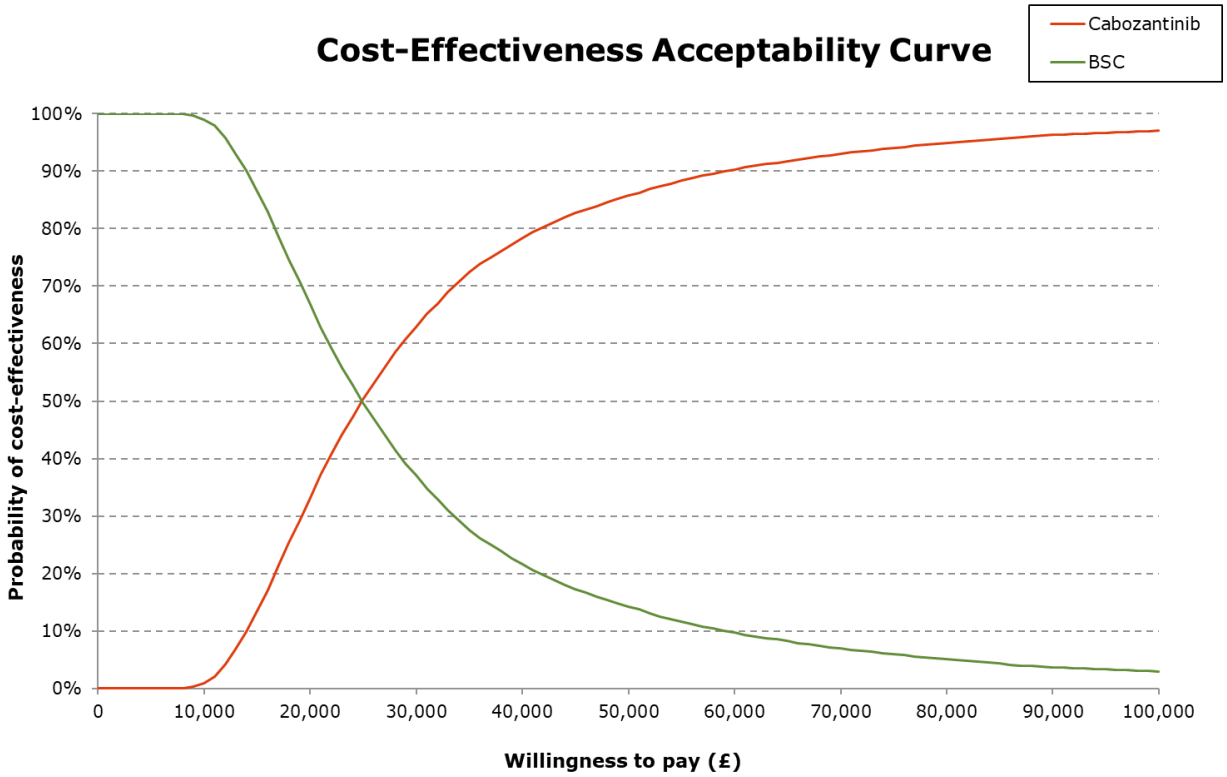
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,867

Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 13: Incremental cost-effectiveness plane - Base case

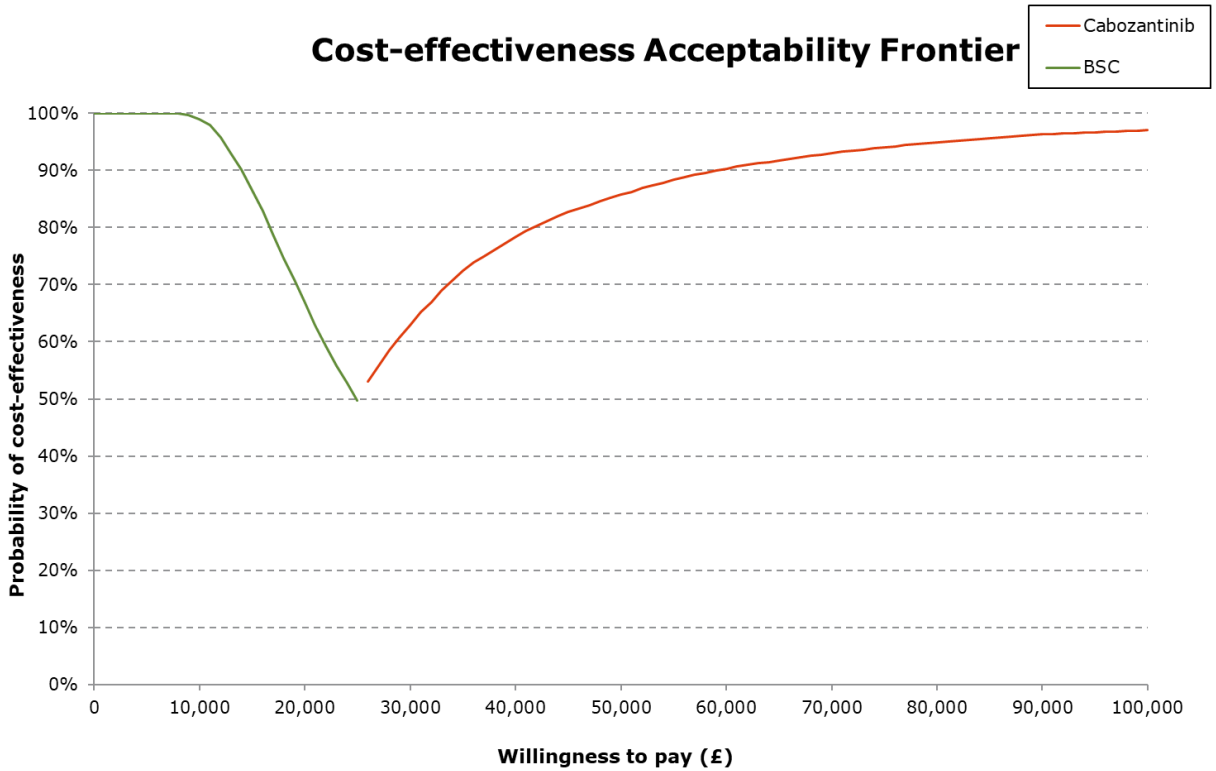
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 14: Cost-effectiveness acceptability curve - Base case



Abbreviations: BSC – Best supportive care

Figure 15: Cost-effectiveness acceptability frontier - Base case



Abbreviations: BSC – Best supportive care

Discount rate – 0%

Table 12: Probabilistic scenario analysis results - Discount rate 0%

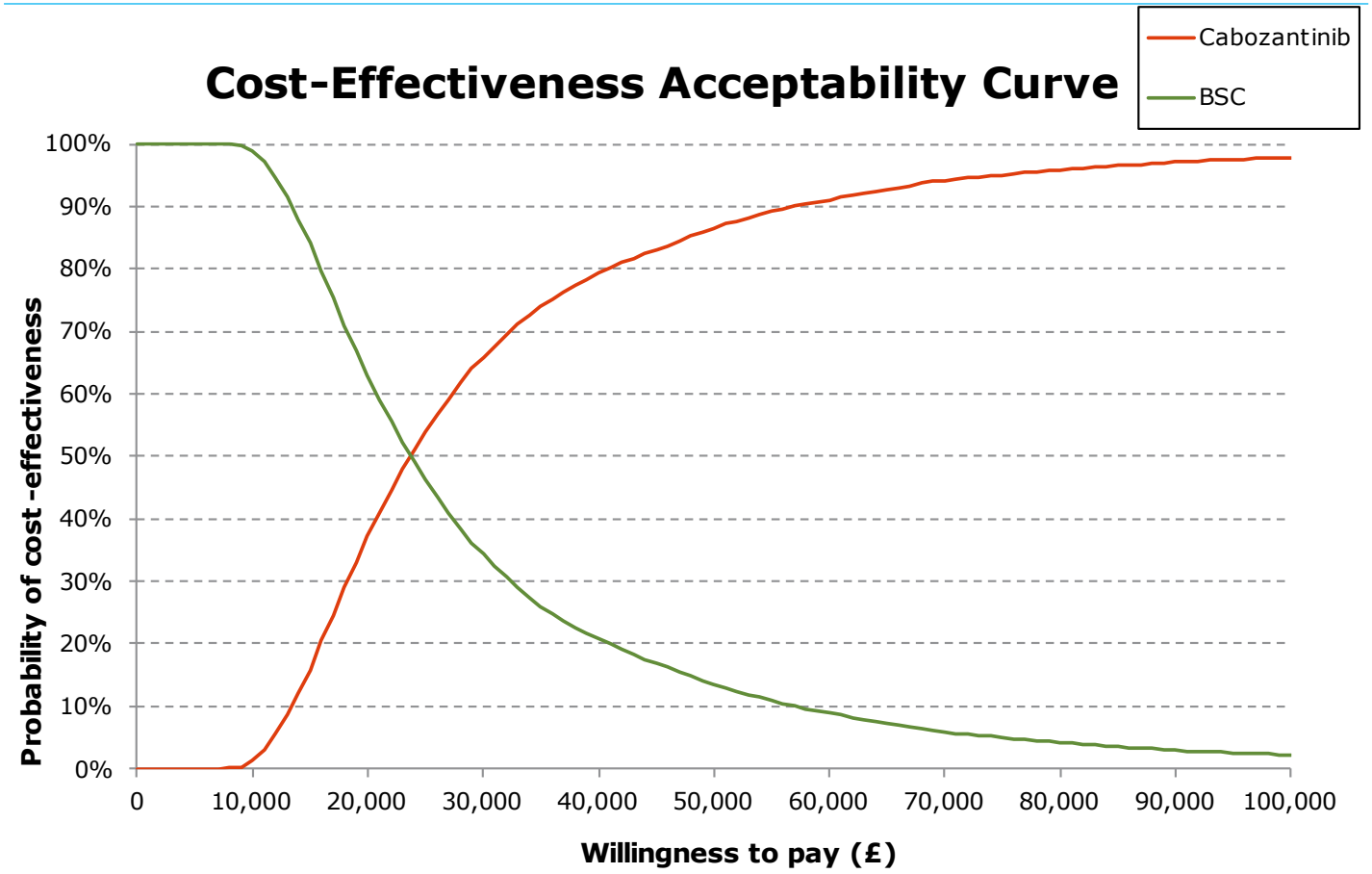
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,815

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 16: Incremental cost-effectiveness plane - Discount 0%

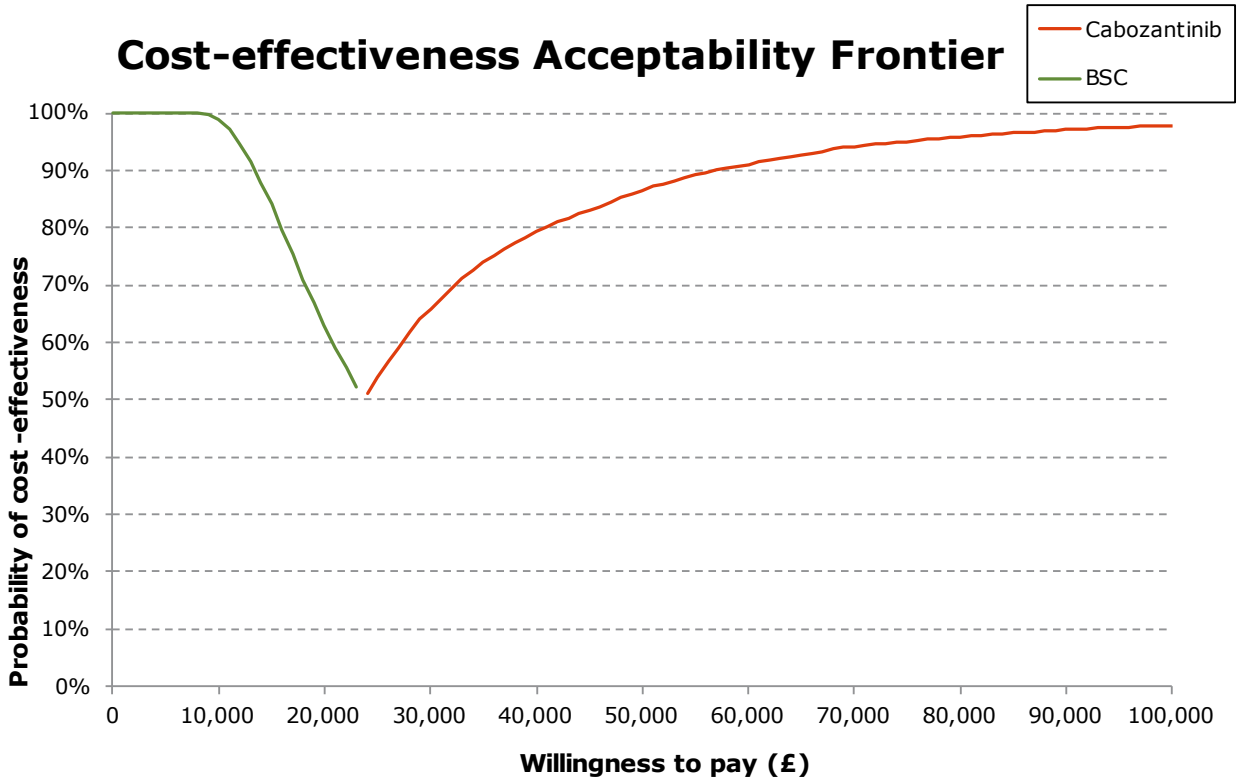
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 17: Cost-effectiveness acceptability curve - Discount 0%



Abbreviations: BSC – Best supportive care

Figure 18: Cost-effectiveness acceptability frontier - Discount 0%



Abbreviations: BSC – Best supportive care

Discount rate – 5%

Table 13: Probabilistic scenario analysis results - Discount rate 5%

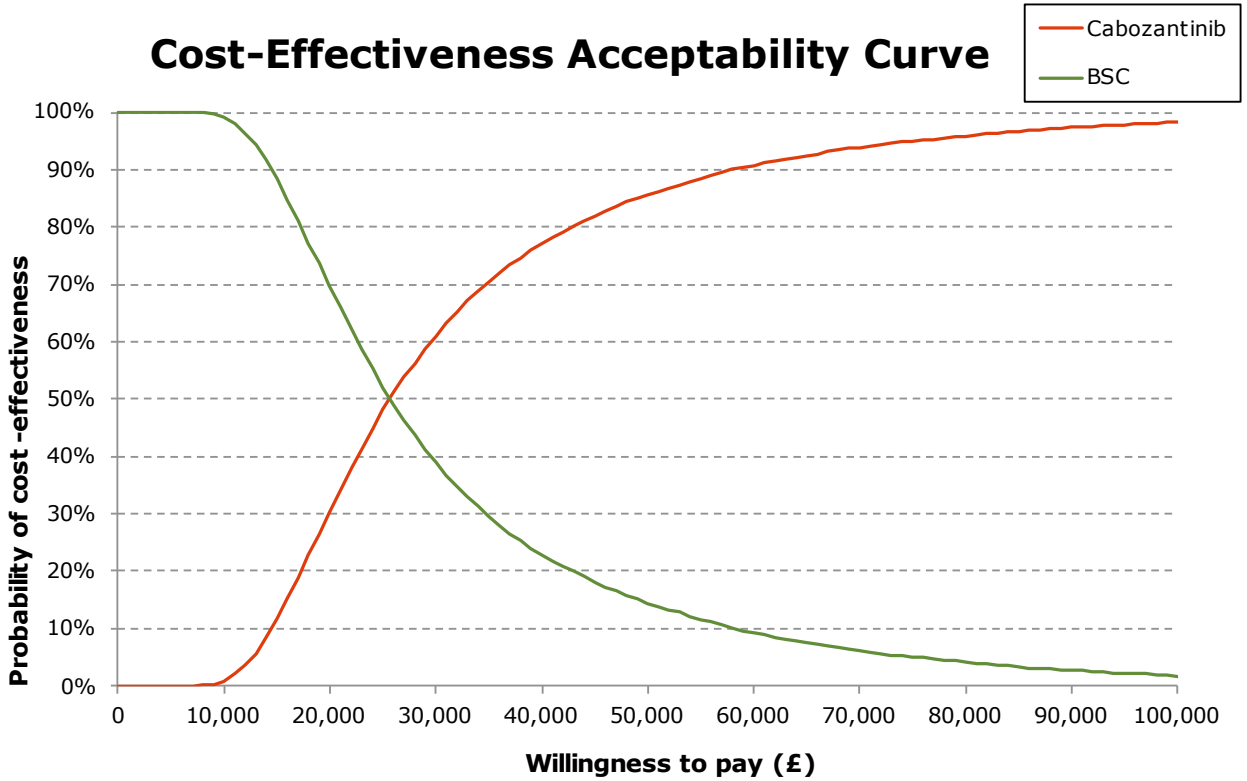
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,440

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 19: Incremental cost-effectiveness plane - Discount 5%

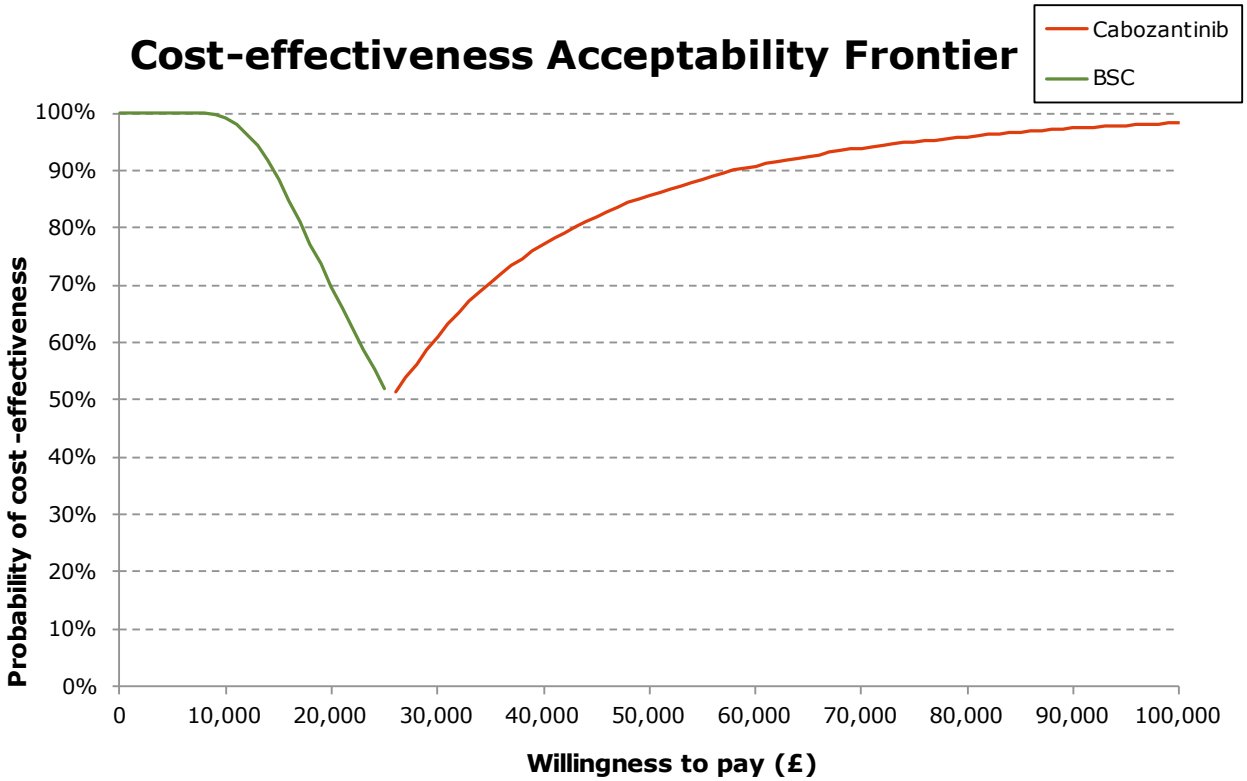
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 20: Cost-effectiveness acceptability curve - Discount 5%



Abbreviations: BSC – Best supportive care

Figure 21: Cost-effectiveness acceptability frontier - Discount 5%



Abbreviations: BSC – Best supportive care

Compliance

Table 14: Probabilistic scenario analysis results – Compliance

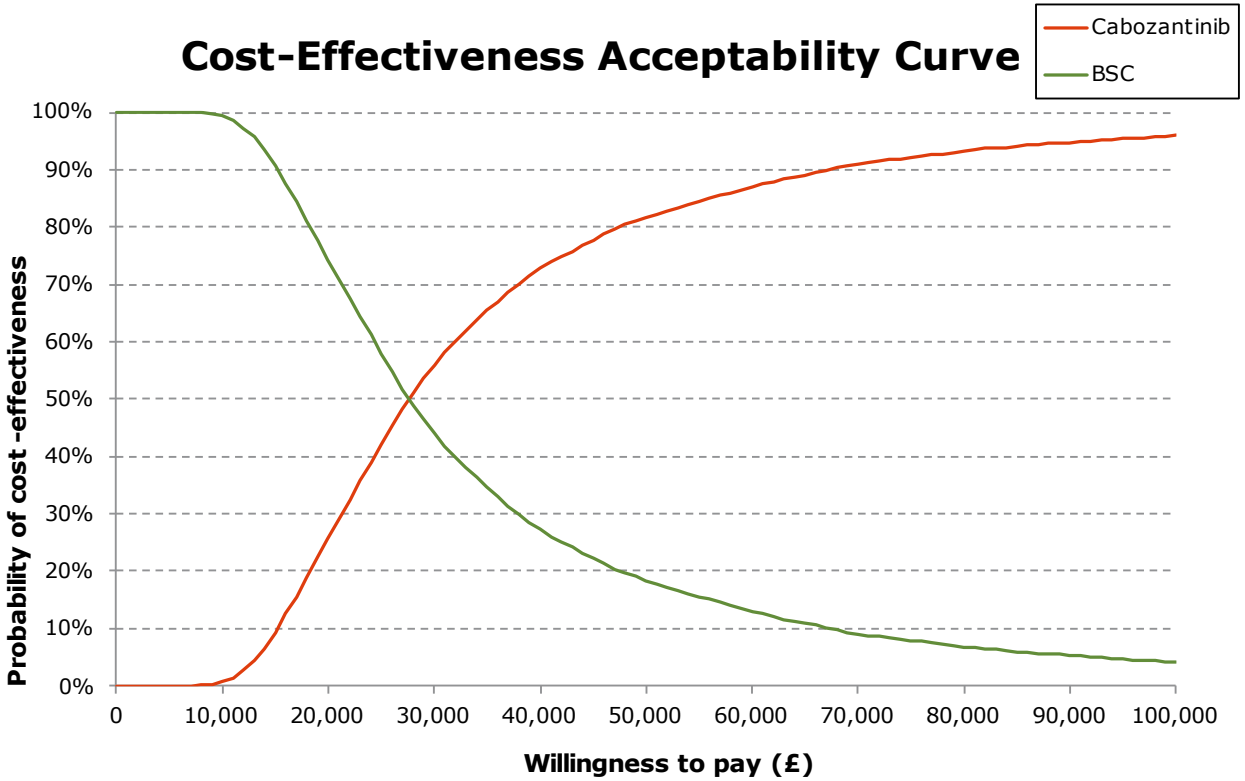
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	23,308

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 22: Incremental cost-effectiveness plane - Compliance

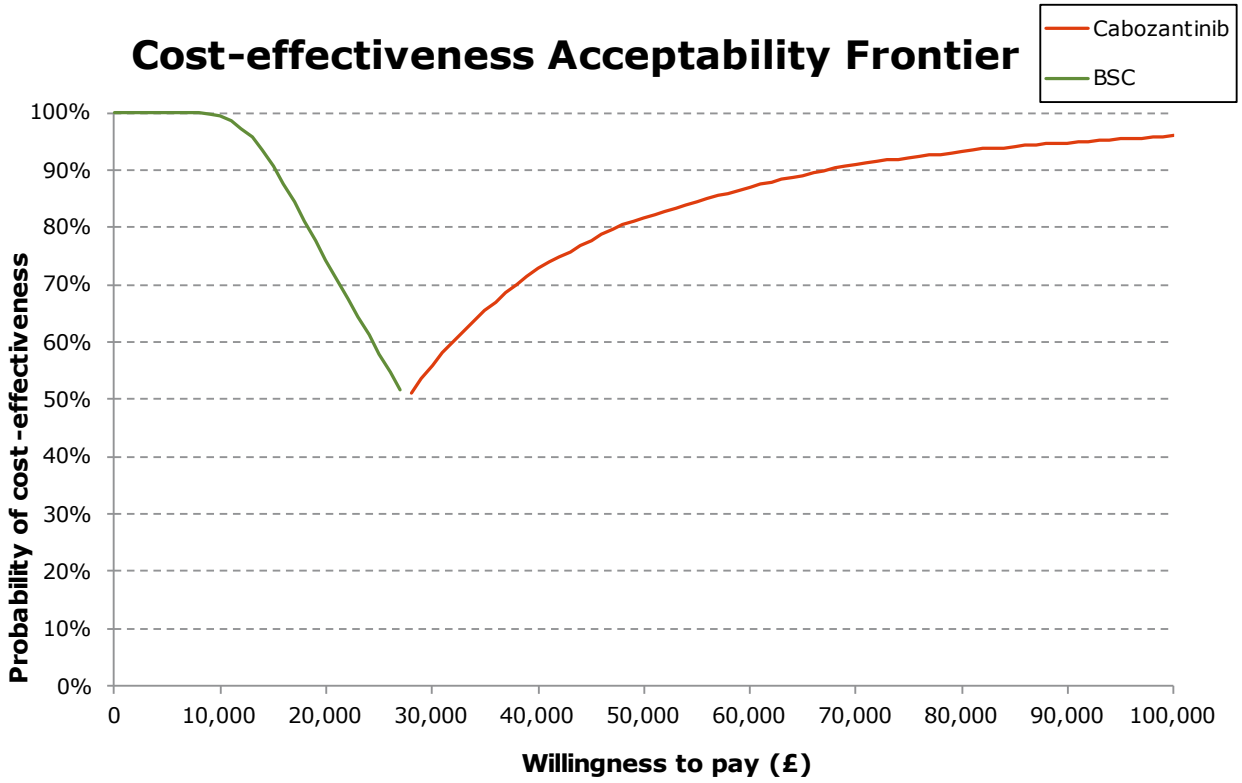
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 23: Cost-effectiveness acceptability curve - Compliance



Abbreviations: BSC – Best supportive care

Figure 24: Cost-effectiveness acceptability frontier - Compliance



Abbreviations: BSC – Best supportive care

5-year OS constraint

Table 15: Probabilistic scenario analysis results – 5-year OS constraint

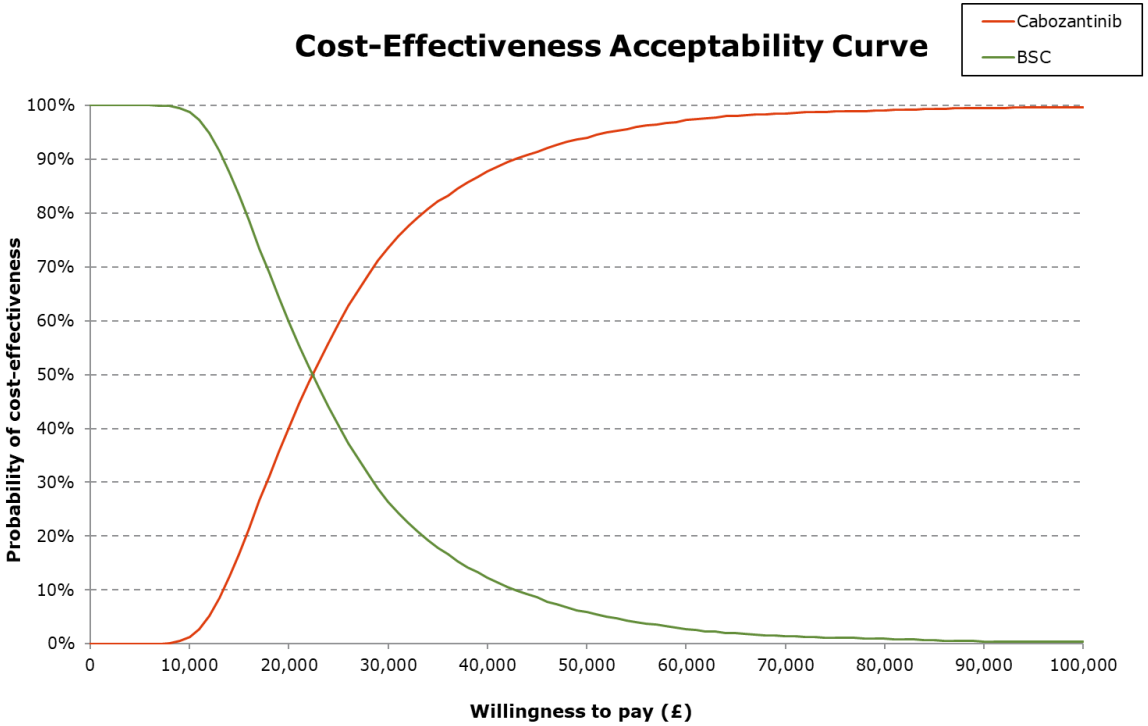
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,015

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 25: Incremental cost-effectiveness plane – 5-year OS constraint

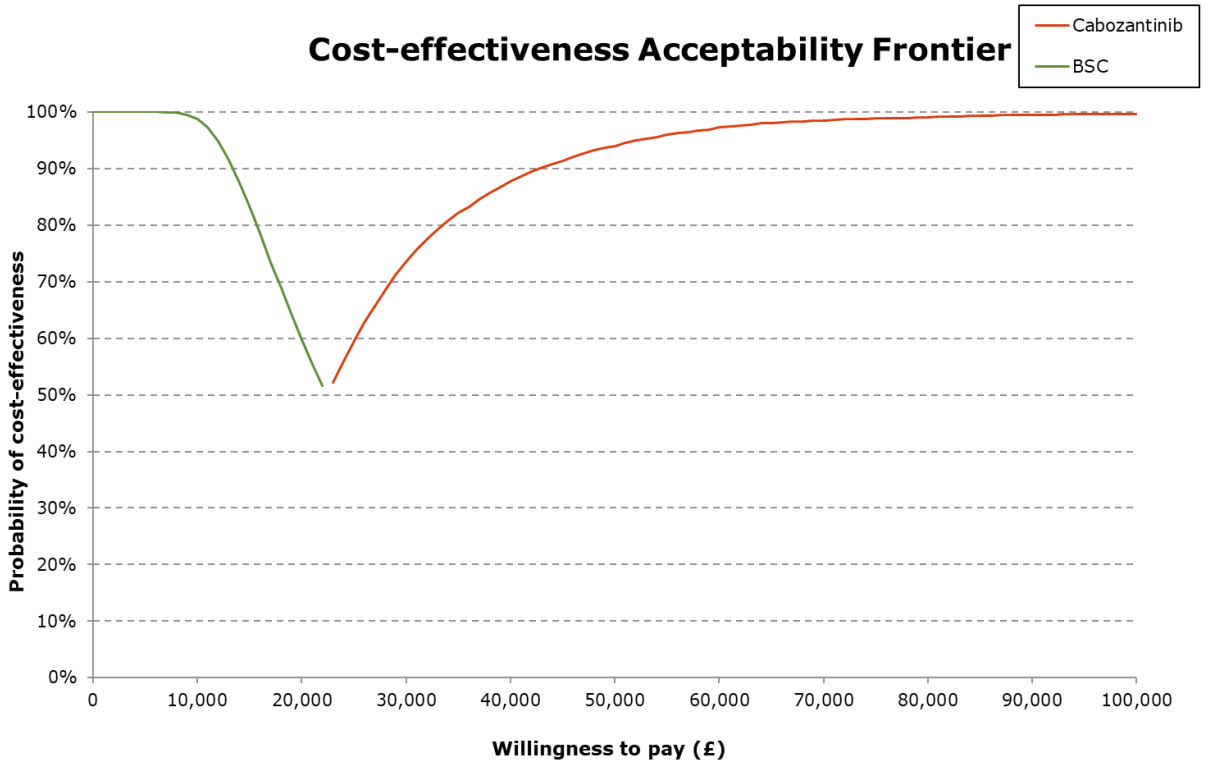
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 26: Cost-effectiveness acceptability curve – 5-year OS constraint



Abbreviations: BSC – Best supportive care

Figure 27: Cost-effectiveness acceptability frontier - 5-year OS constraint



Abbreviations: BSC – Best supportive care

OS curve: lognormal

Table 16: Probabilistic scenario analysis results – OS curve: lognormal

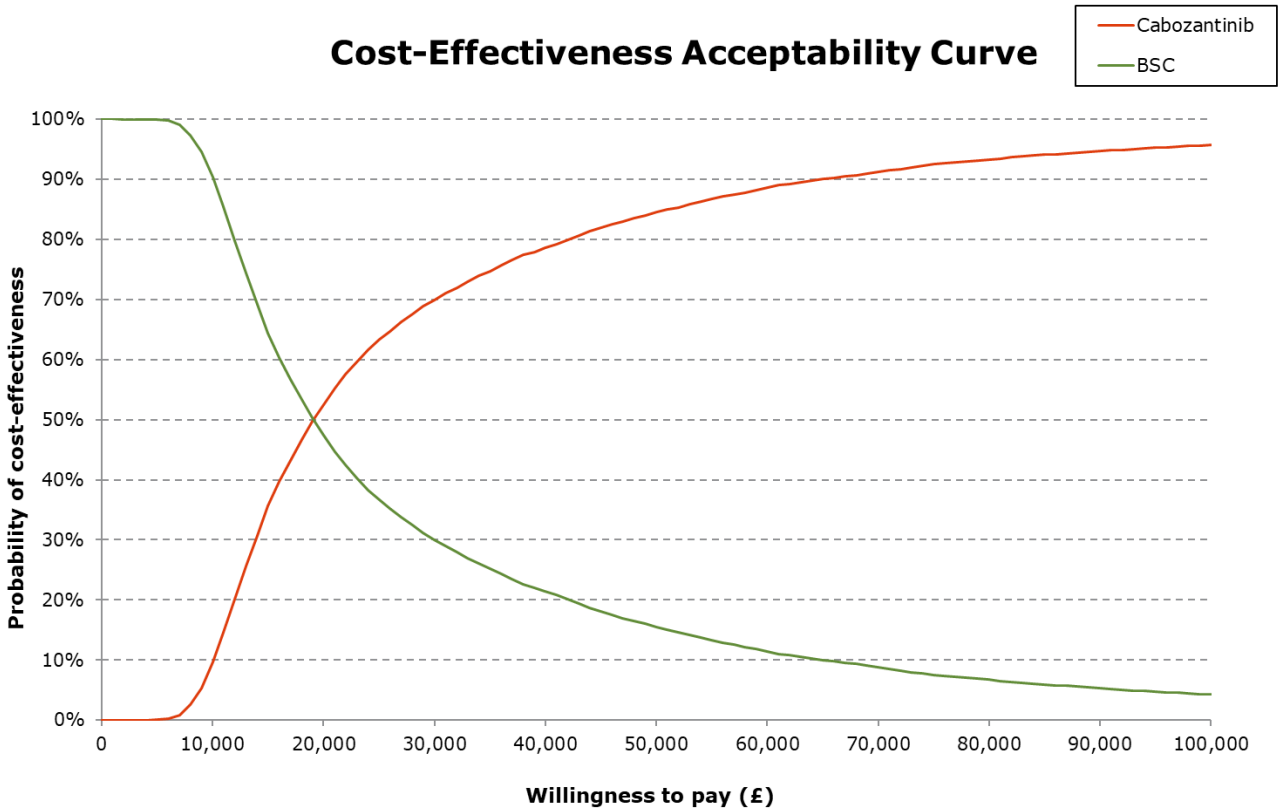
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	15,049

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 28: Incremental cost-effectiveness plane – OS curve: lognormal

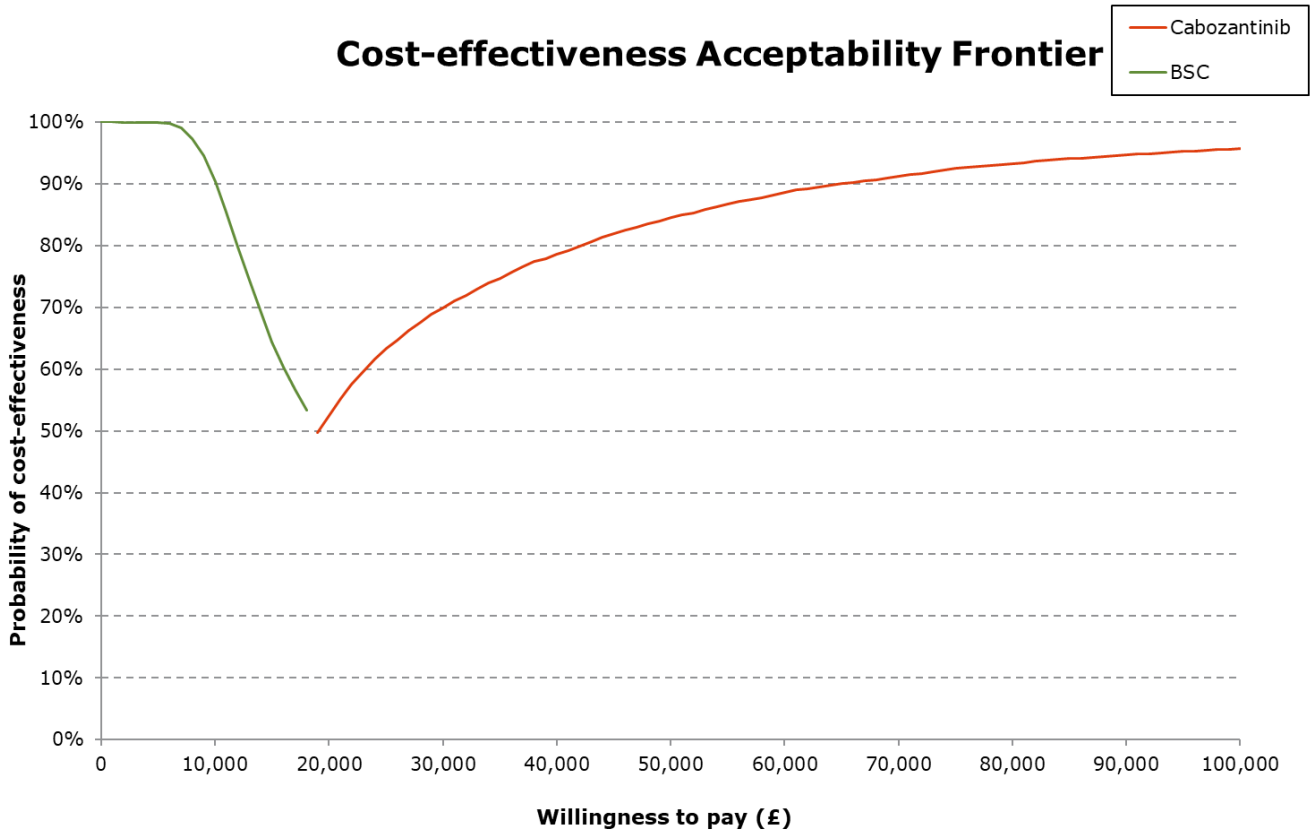
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 29: Cost-effectiveness acceptability curve – OS curve: lognormal



Abbreviations: BSC – Best supportive care

Figure 30: Cost-effectiveness acceptability frontier – OS curve: lognormal



Abbreviations: BSC – Best supportive care

OS curve: log-logistic

Table 17: Probabilistic scenario analysis results – OS curve: log-logistic

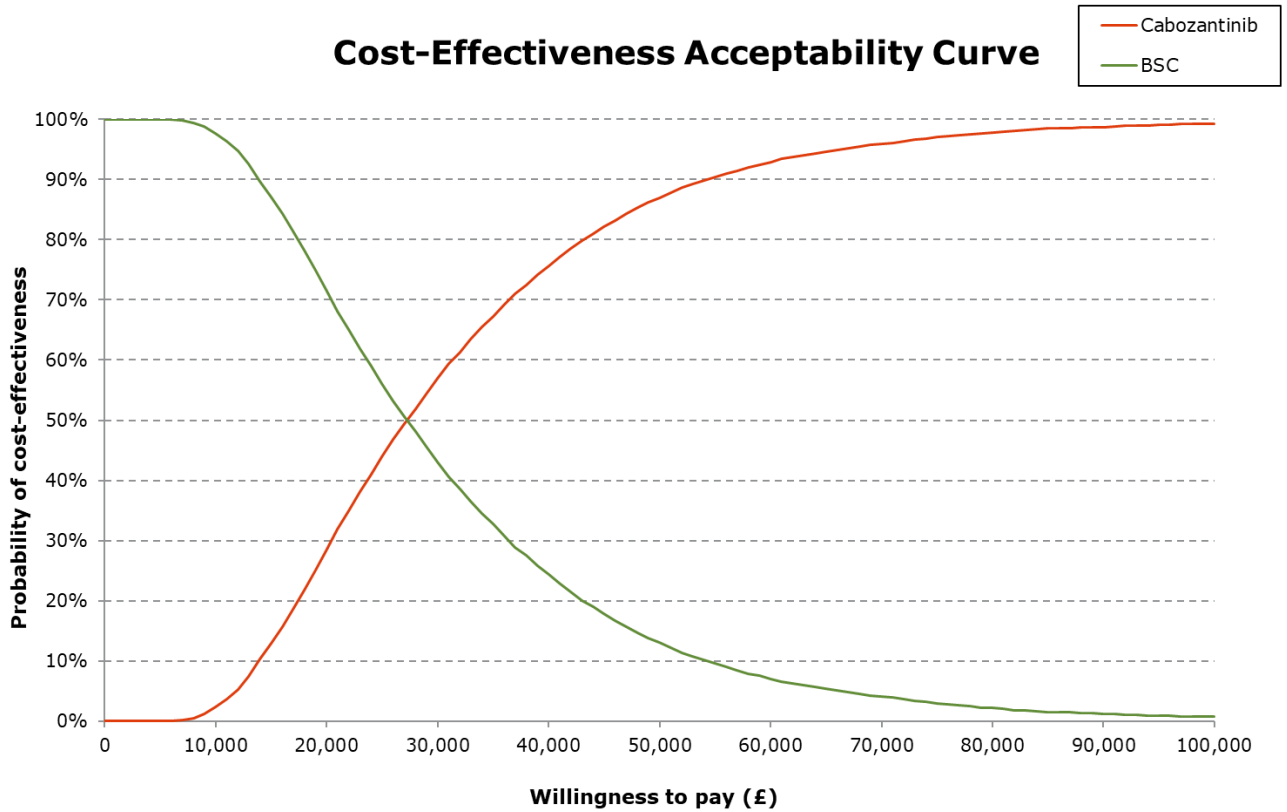
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,221

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years ; QALYs – Quality-adjusted life years

Figure 31: Incremental cost-effectiveness plane – OS curve: log-logistic

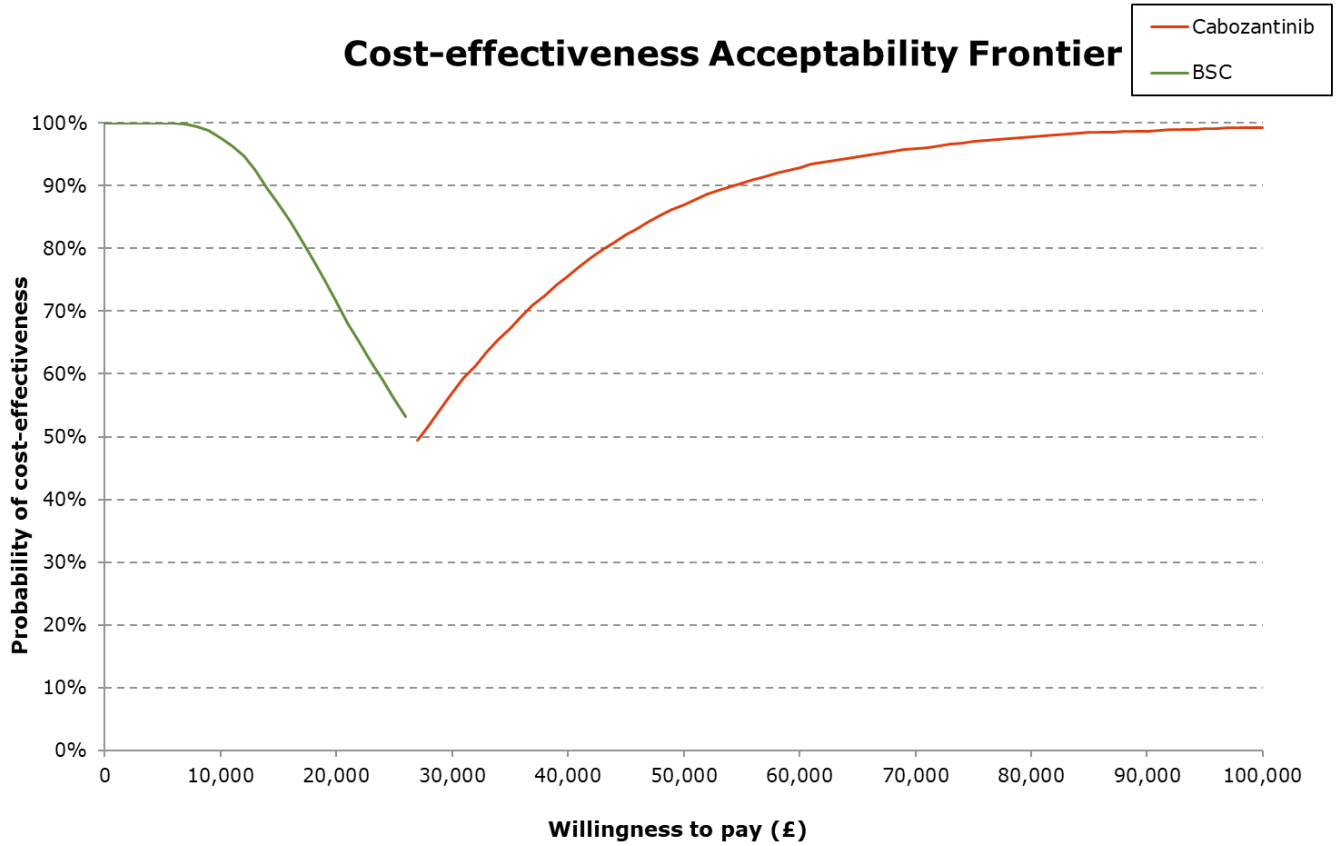
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 32: Cost-effectiveness acceptability curve – OS curve: log-logistic



Abbreviations: BSC – Best supportive care

Figure 33: Cost-effectiveness acceptability frontier – OS curve: log-logistic



Abbreviations: BSC – Best supportive care

PFS curve: lognormal

Table 18: Probabilistic scenario analysis results – PFS curve: lognormal

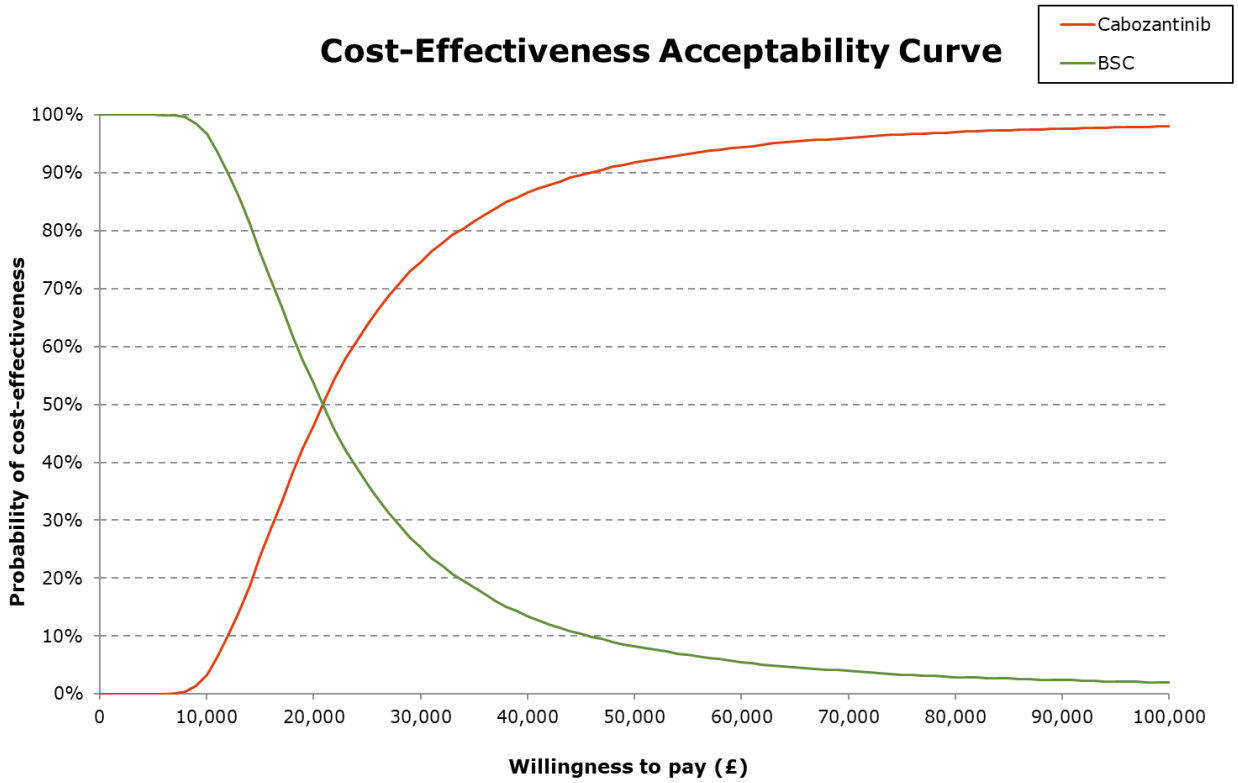
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,474

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 34: Incremental cost-effectiveness plane – PFS curve: lognormal

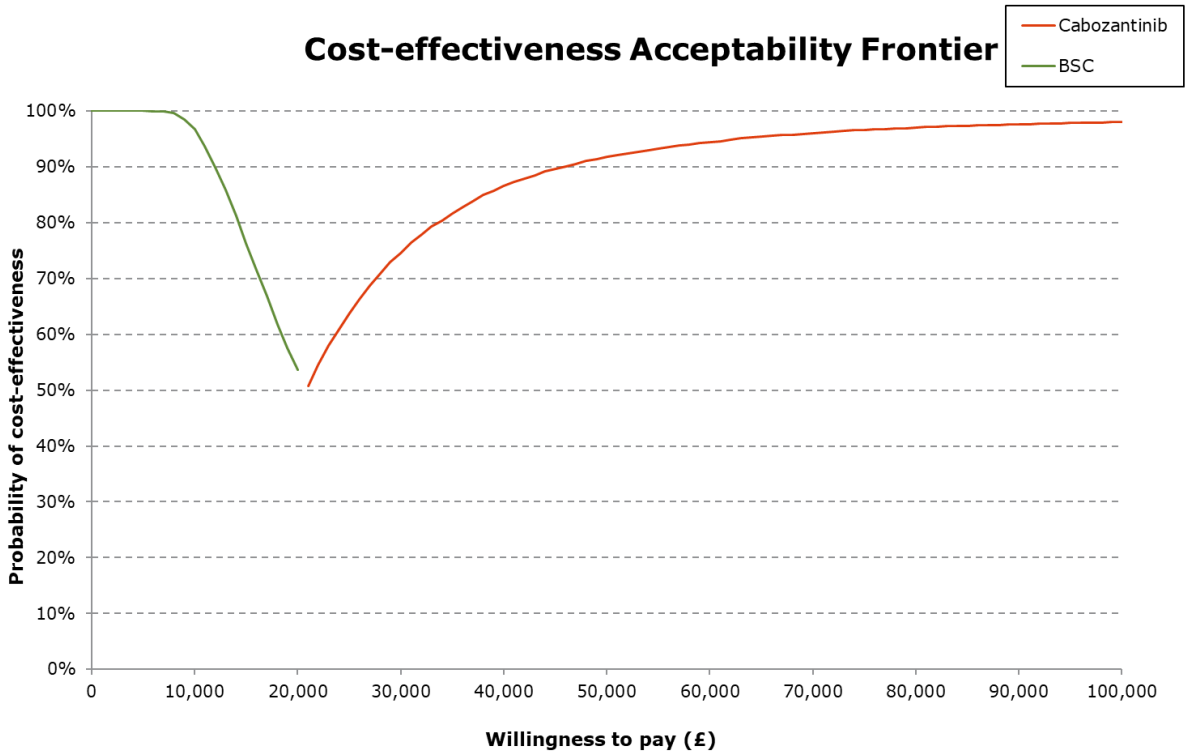
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 35: Cost-effectiveness acceptability curve – PFS curve: lognormal



Abbreviations: BSC – Best supportive care

Figure 36: Cost-effectiveness acceptability frontier – PFS curve: lognormal



Abbreviations: BSC – Best supportive care

PFS curve: generalised gamma

Table 19: Probabilistic scenario analysis results - PFS curve: generalised gamma

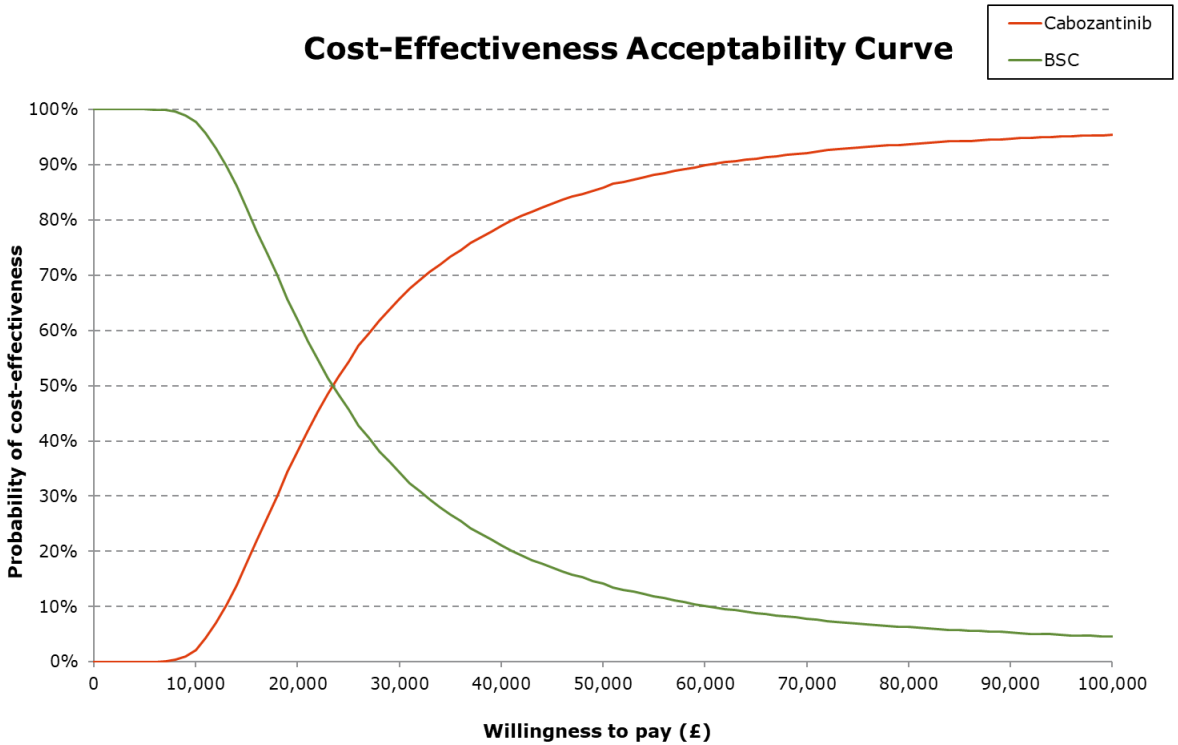
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,552

*Severity modifier of 1.2 has been applied to incremental QALYs
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 37: Incremental cost-effectiveness plane – PFS curve: generalised gamma

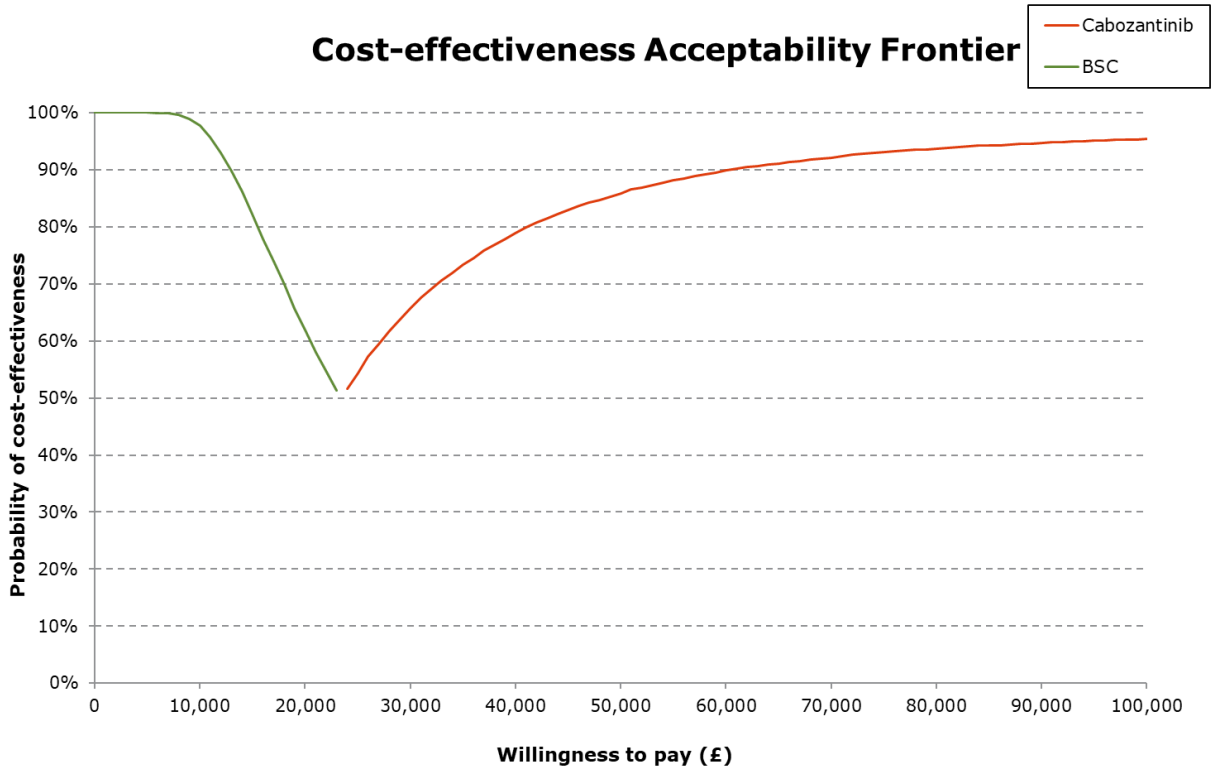
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 38: Cost-effectiveness acceptability curve – PFS curve: generalised gamma



Abbreviations: BSC – Best supportive care

Figure 39: Cost-effectiveness acceptability frontier – PFS curve: generalised gamma



Abbreviations: BSC – Best supportive care

TTD curve: exponential

Table 20: Probabilistic scenario analysis results – TTD curve: exponential

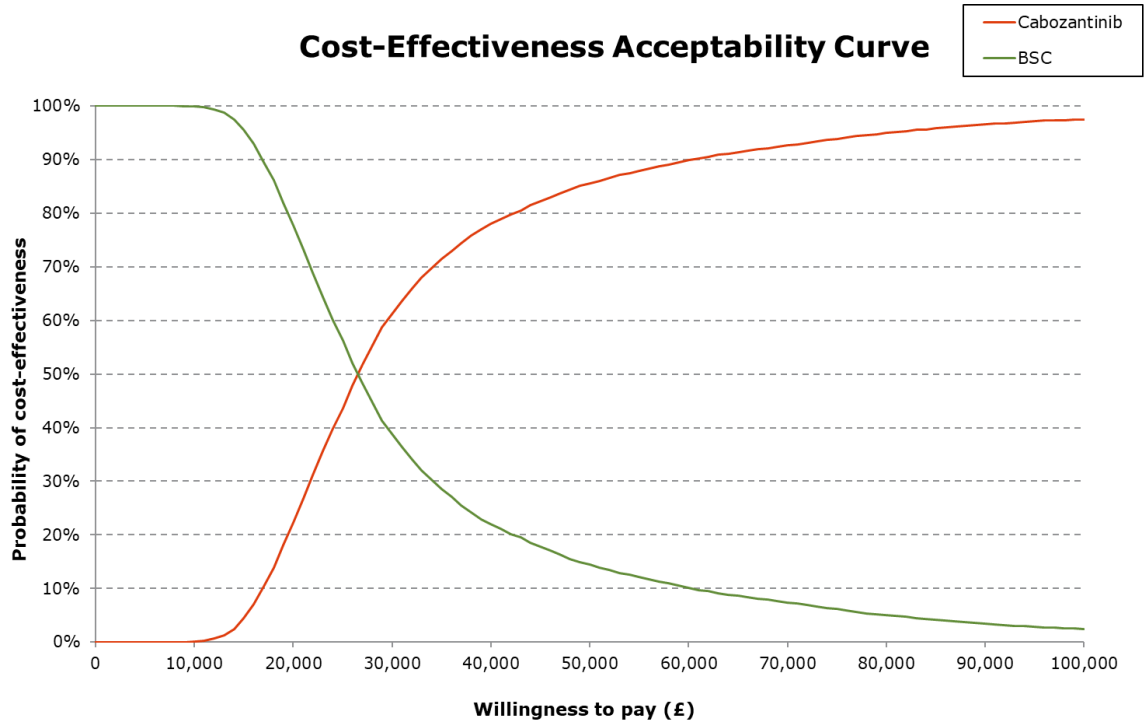
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,336

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 40: Incremental cost-effectiveness plane - TTD curve: exponential

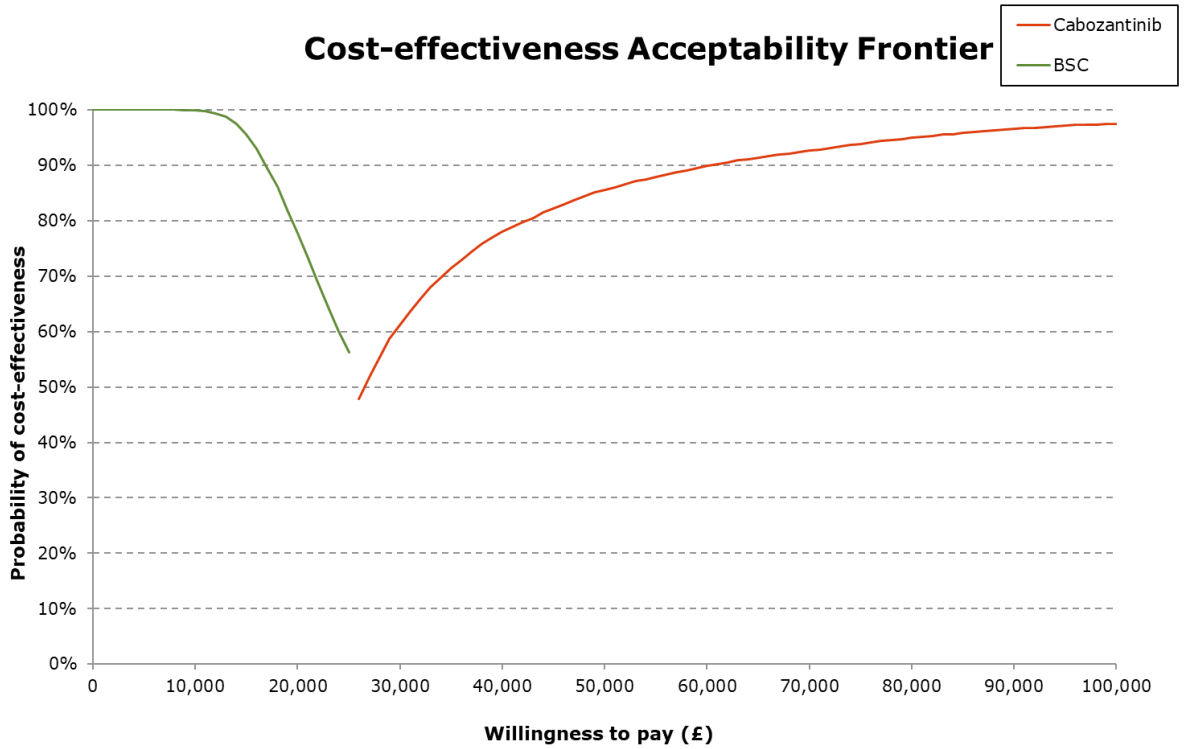
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 41: Cost-effectiveness acceptability curve - TTD curve: exponential



Abbreviations: BSC – Best supportive care

Figure 42: Cost-effectiveness acceptability frontier - TTD curve: exponential



Abbreviations: BSC – Best supportive care

TTD curve: Gompertz

Table 21: Probabilistic scenario analysis results – TTD curve: Gompertz

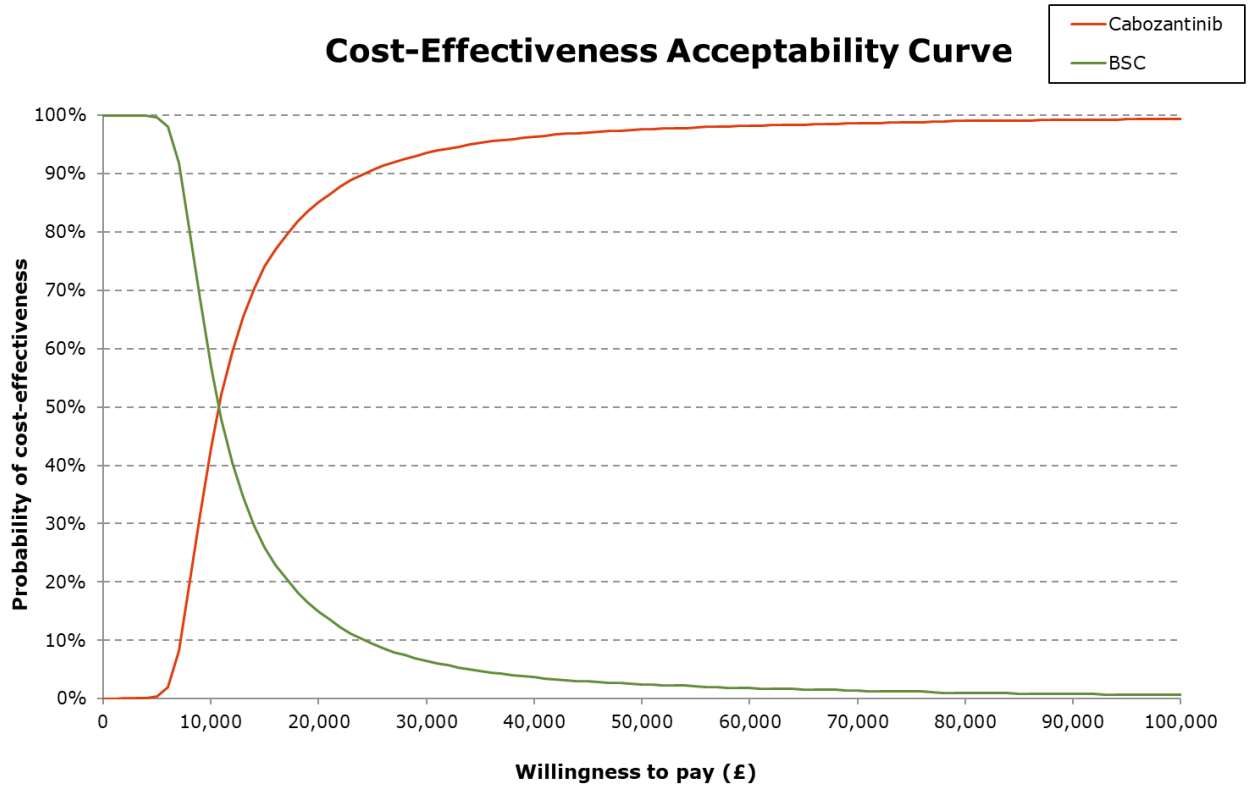
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	10,122

**Severity modifier of 1.2 has been applied to incremental QALYs*
Abbreviations: BSC – Best supportive care; ICER – Incremental cost-effectiveness ratio; LYG – Life years gained; QALYs – Quality-adjusted life years

Figure 43: Incremental cost-effectiveness plane – TTD curve: Gompertz

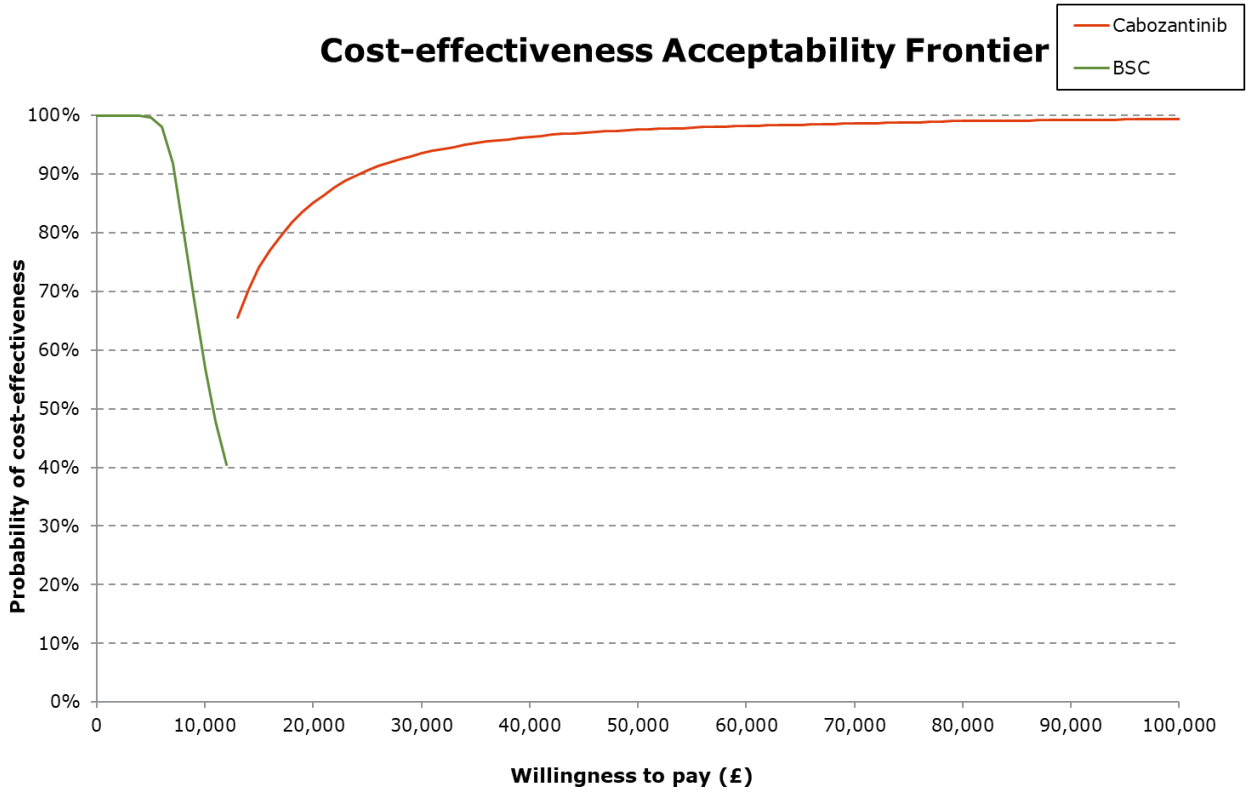
Abbreviations: BSC – Best supportive care; PSA – Probabilistic sensitivity analysis

Figure 44: Cost-effectiveness acceptability curve – TTD curve: Gompertz



Abbreviations: BSC – Best supportive care

Figure 45: Cost-effectiveness acceptability frontier - TTD curve: Gompertz



Abbreviations: BSC – Best supportive care

Blended survival analysis results

Ipsen conducted a blended survival analysis incorporating clinical expert opinion into CCO2 data. To reduce uncertainty around long-term OS estimates for cabozantinib and BSC, the clinical expert opinion was based on the mean average across EAG and company advisors as per Table 1.

Scenario – blending base case

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,521

Figure 46: Incremental cost-effectiveness plane – Blended survival base case scenario

Figure 47: Cost-effectiveness acceptability curve – Blended survival base case scenario

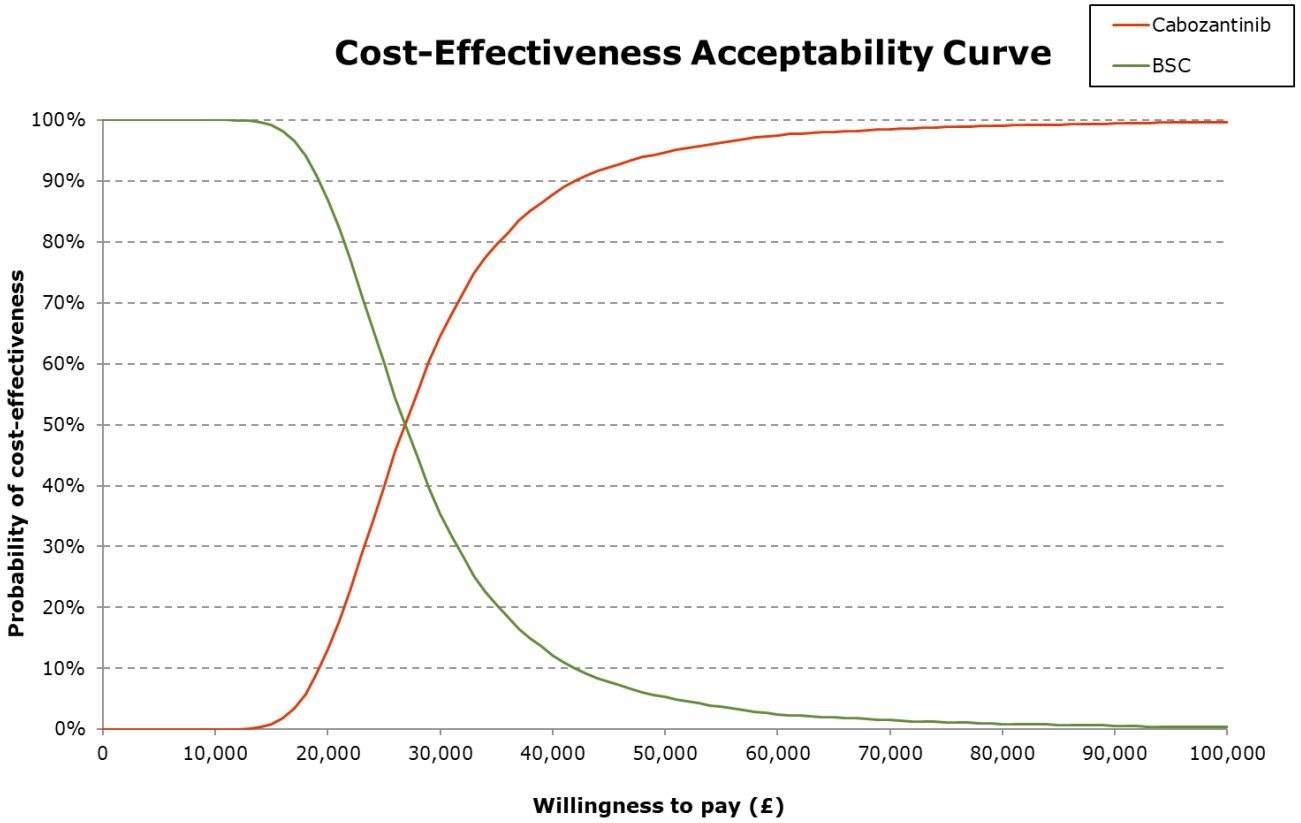
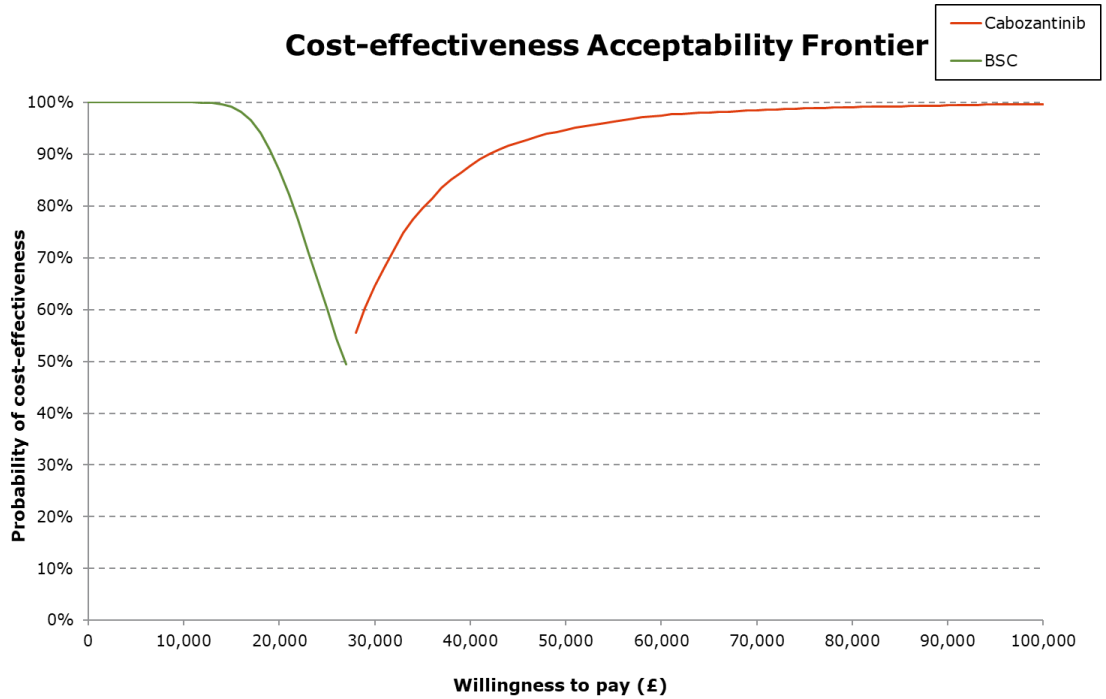


Figure 48: Cost-effectiveness acceptability frontier– Blended survival base case scenario



Scenario – blending interval (60 months)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)

BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,538

Figure 49: Incremental cost-effectiveness plane – Blended survival blending interval (60 months)

Figure 50: Cost-effectiveness acceptability curve – Blended survival blending interval (60 months)

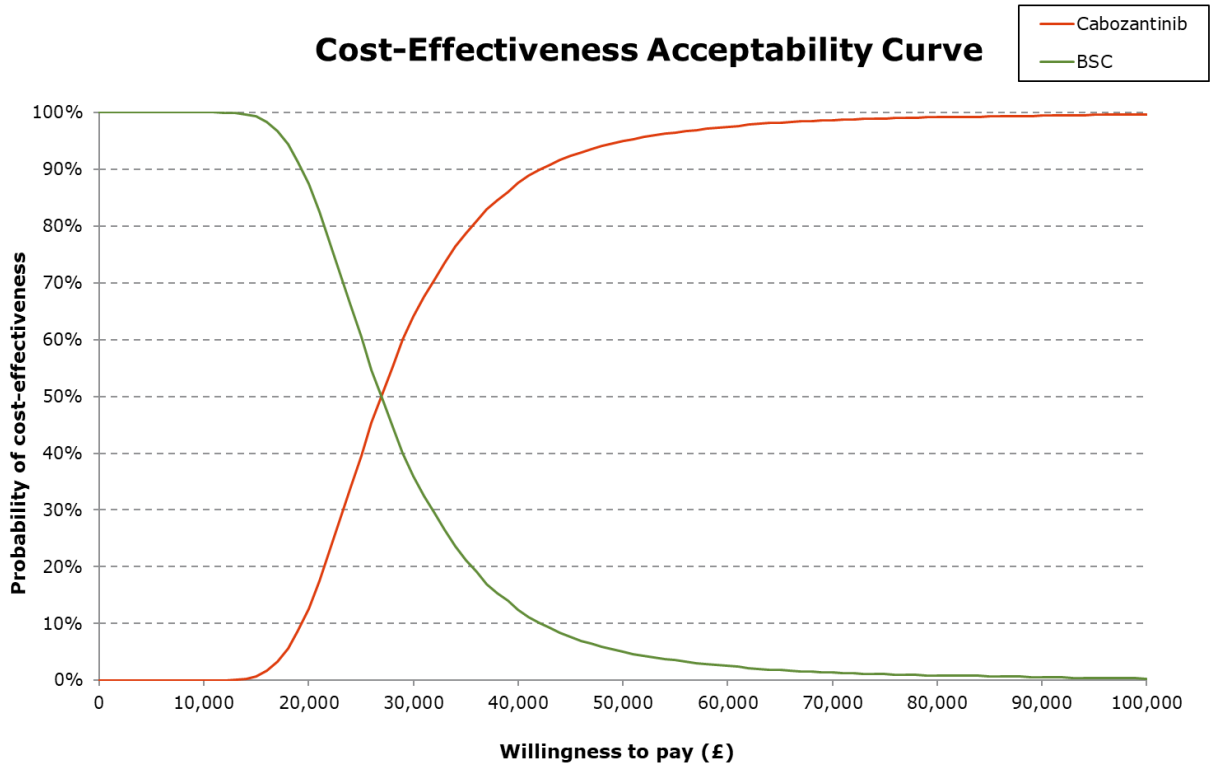
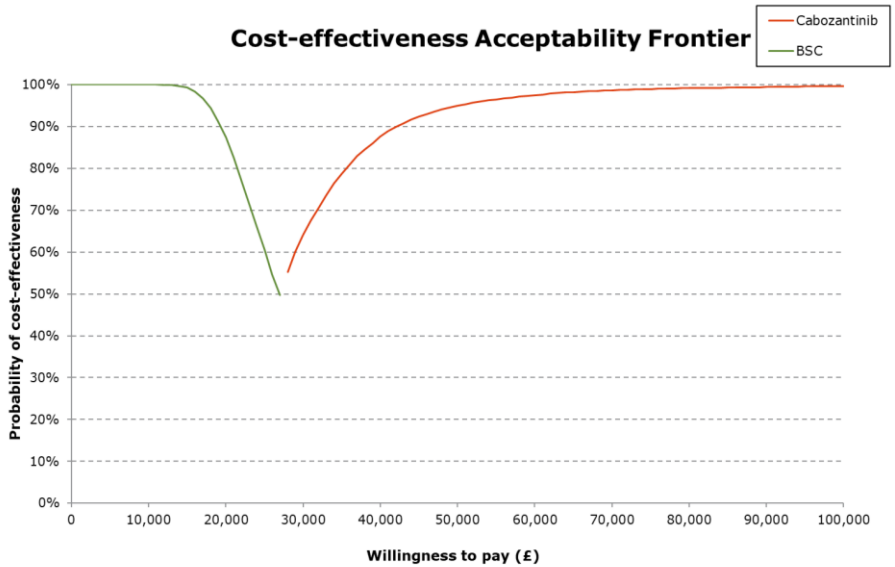


Figure 51: Cost-effectiveness acceptability frontier– Blended survival blending interval (60 months)



Scenario – blending rate (Alpha=2, Beta=5)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,475

Figure 52: Incremental cost-effectiveness plane – Blended survival blending rate (Alpha=2, Beta=5)

Figure 53: Cost-effectiveness acceptability curve – Blended survival blending rate (Alpha=2, Beta=5)

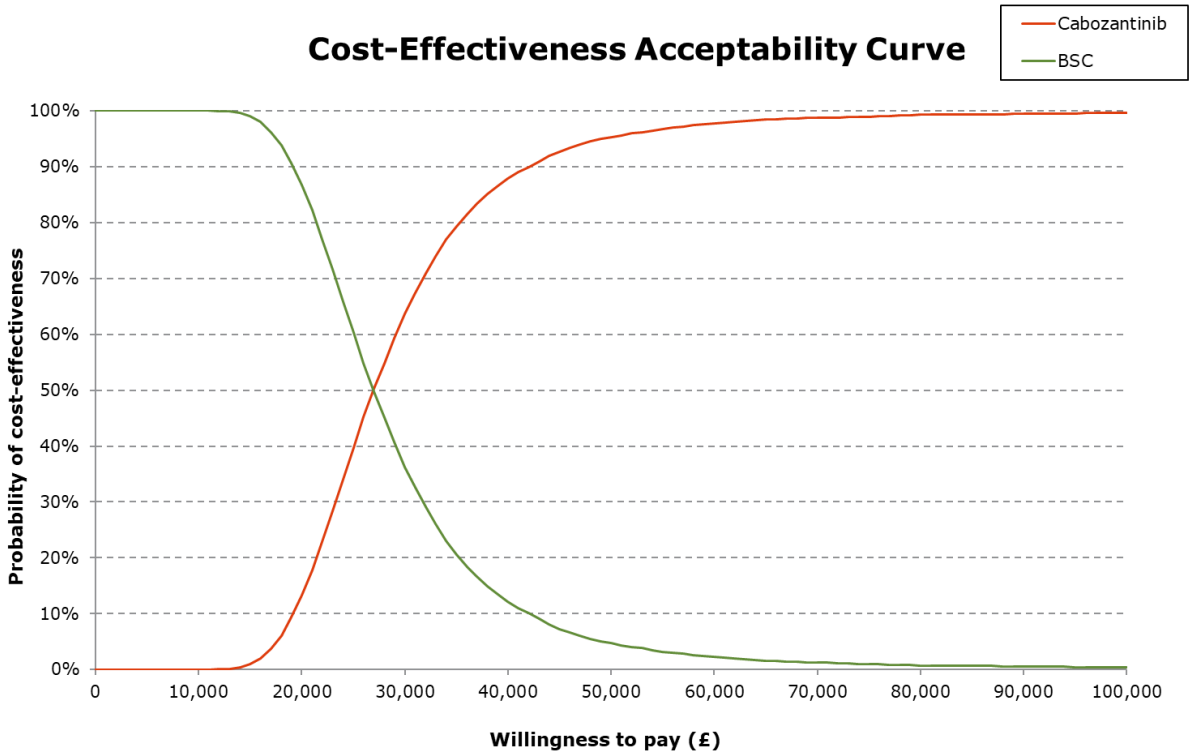
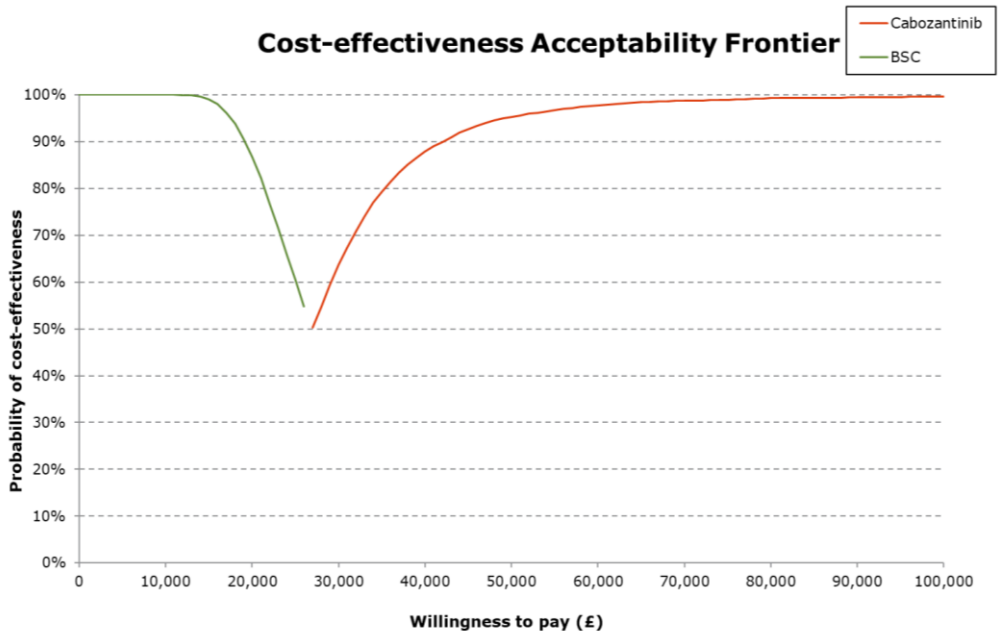


Figure 54: Cost-effectiveness acceptability frontier – Blended survival blending rate (Alpha=2, Beta=5)



Scenario – high uncertainty

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	23,552

Figure 55: Incremental cost-effectiveness plane – Blended survival high uncertainty

[redacted]

Figure 56: Cost-effectiveness acceptability curve – Blended survival high uncertainty

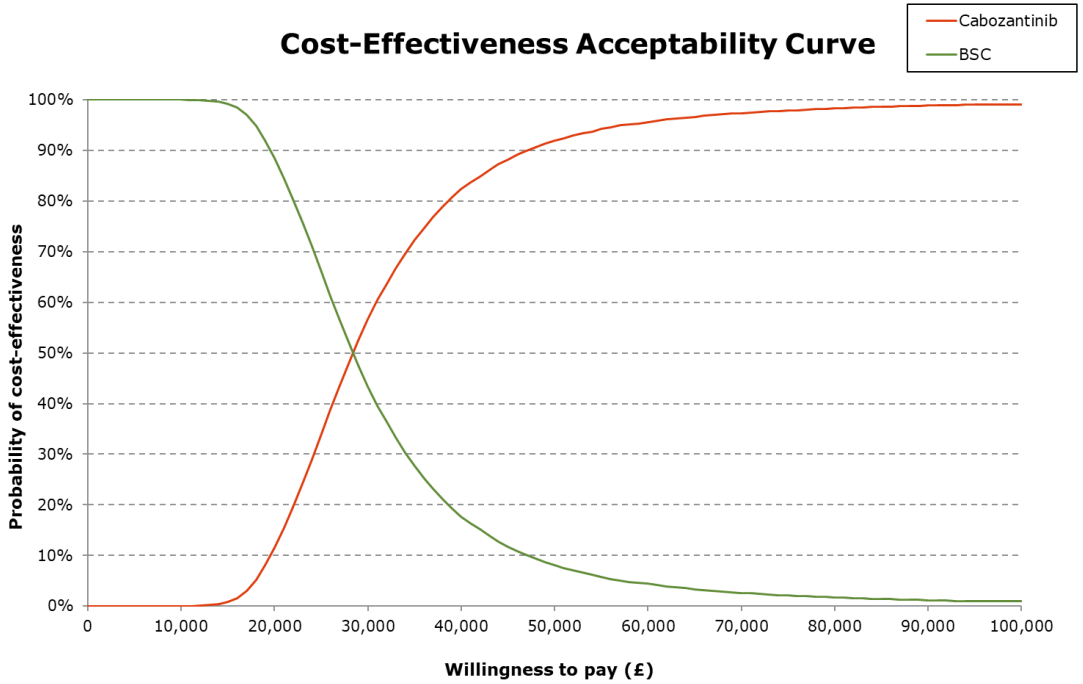
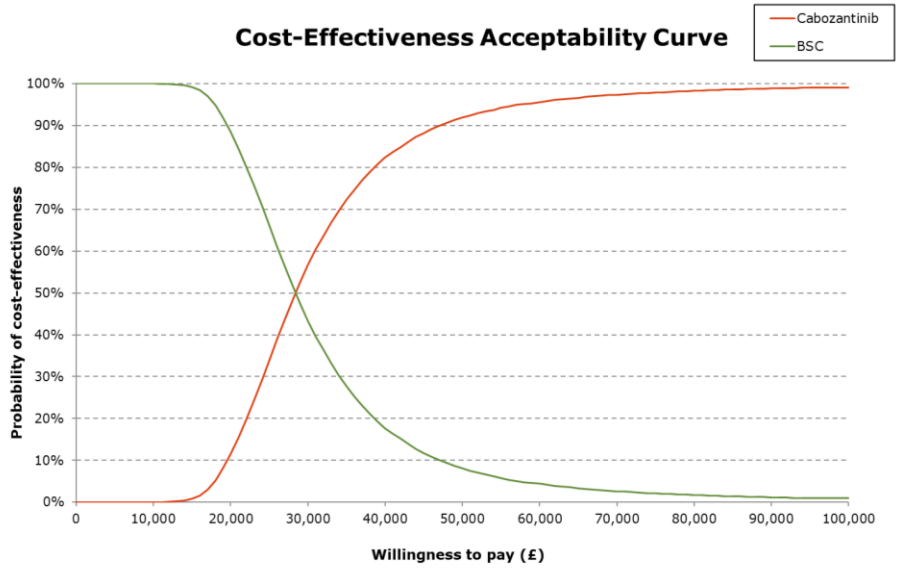


Figure 57: Cost-effectiveness acceptability frontier– Blended survival high uncertainty



Scenario – low uncertainty

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Severity ICER incremental* (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,208

Figure 58: Incremental cost-effectiveness plane – Blended survival low uncertainty

[redacted]

Figure 59: Cost-effectiveness acceptability curve – Blended survival low uncertainty

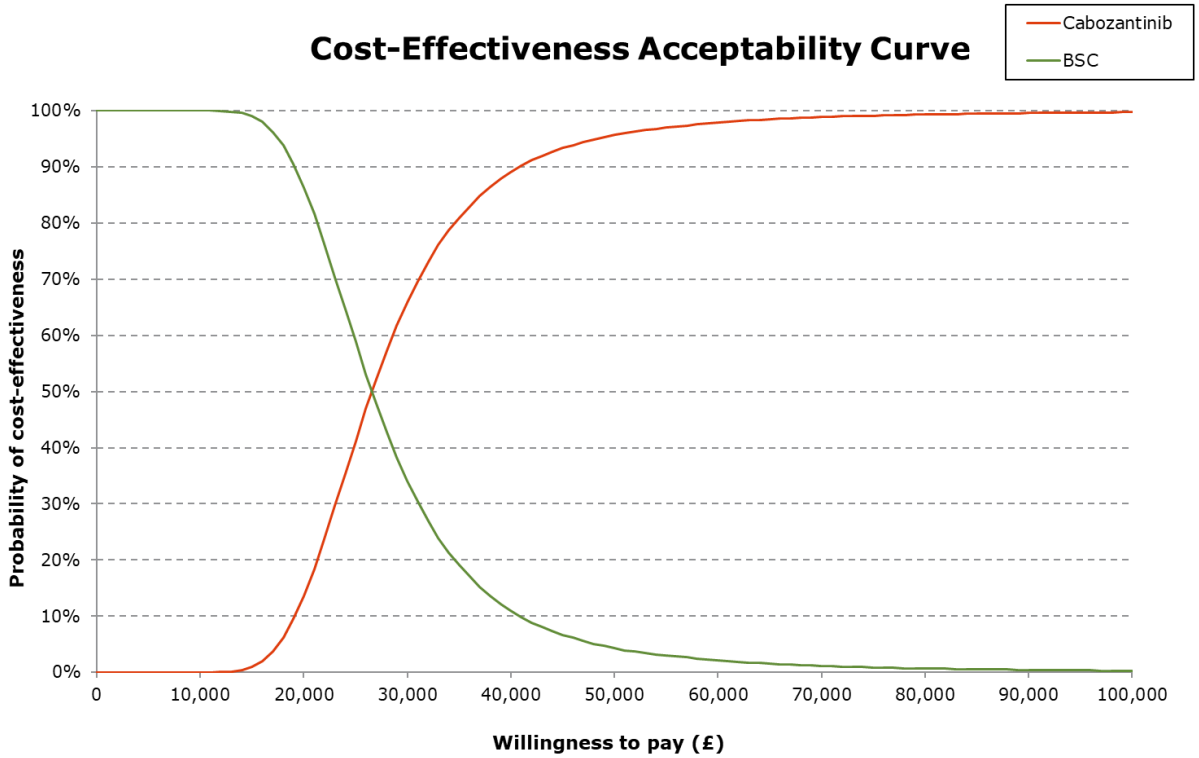
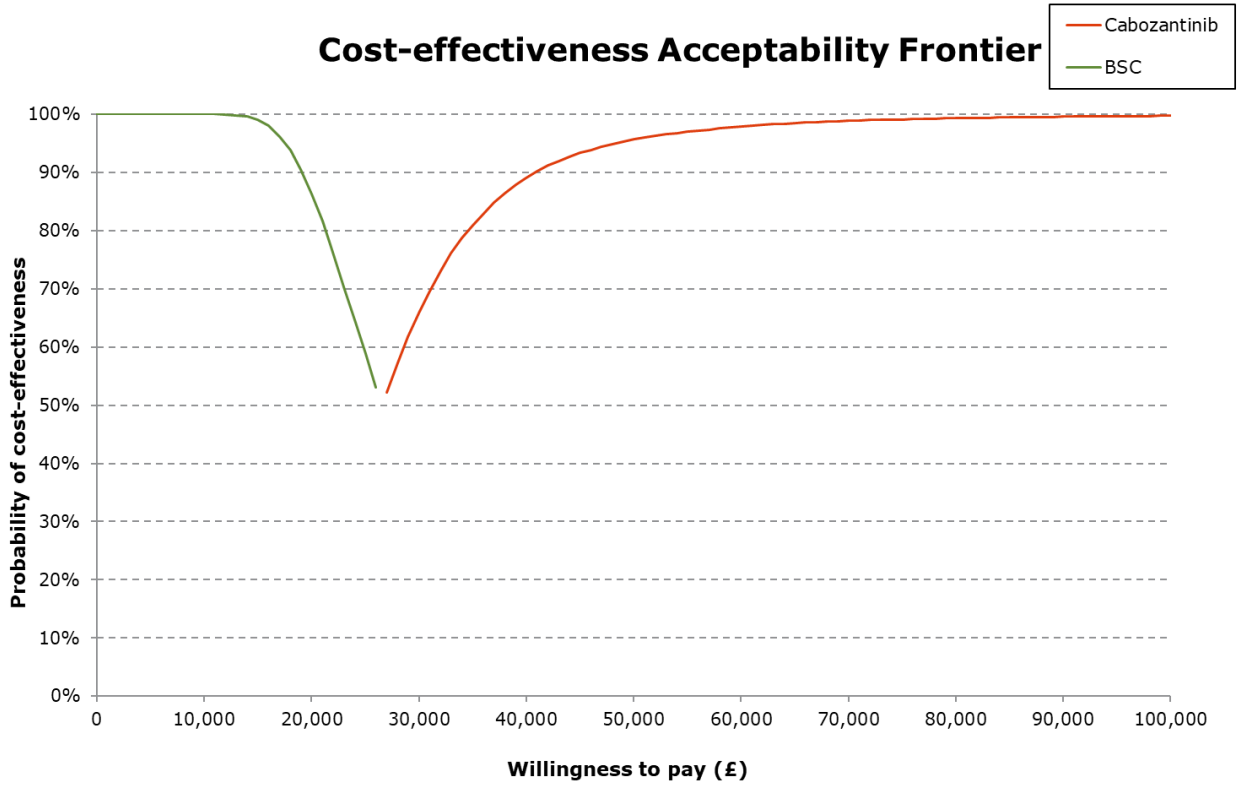


Figure 60: Cost-effectiveness acceptability frontier– Blended survival low uncertainty



NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Clinical expert statement and technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing 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. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

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. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **Wednesday 25th January 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Treating thyroid cancer and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Dr Sath Nag
2. Name of organisation	South Tees Hospitals NHS Foundation Trust
3. Job title or position	Consultant Endocrinologist
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 thyroid cancer? ☐ A specialist in the clinical evidence base for thyroid cancer 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 did not submit one, I do not 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
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	No declarations
8. What is the main aim of treatment for differentiated thyroid cancer which is unsuitable for or refractory to	Improve survival of patients without disabling tumour burden, improve quality of life and minimise discomfort due to progressive metastases, arrest disease progression and reduce disability due to residual tumour burden.

radioactive iodine and has progressed during or after prior systemic therapy? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	
9. 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)	Reduction in tumour burden of more than 70%
10. In your view, is there an unmet need for patients and healthcare professionals in differentiated thyroid cancer which is unsuitable for or refractory to radioactive iodine and has progressed during or after prior systemic therapy?	There is an unmet need for patients with metastatic differentiated thyroid cancer whose disease becomes refractory or unresponsive to further treatment with radioactive iodine or has progressed following treatment with other Tyrosine Kinase inhibitor drugs. Differentiated thyroid cancer affects people of working age and refractory disease poses huge challenges in terms of impacting on patient's ability to lead normal and productive lives. Given that, in general, differentiated thyroid cancer has a relatively good long-term prognosis, efforts should be focussed in improving the quality of lives with patients with radio-refractory disease. This serves to reduce long term disability, reduce tumour burden, and increase overall survival.
11. How is differentiated thyroid cancer which is unsuitable for or refractory to radioactive iodine and has progressed during or after prior systemic therapy currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals	Most Thyroid cancer MDT's in the UK work to thyroid cancer guidelines produced by the British Thyroid Association (2014) and the American Thyroid Association (2016). NICE Technology appraisals also inform treatment strategies in patients with differentiated thyroid cancer who 'escape' the effects of treatment dose radioactive iodine or who become refractory or unresponsive to further treatments with radio-iodine due to de-differentiation of thyroid cancers.

across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	The current pathway of care for thyroid cancer patients in the UK is well defined and standardised as all clinicians practice in accordance with NICE guidance and use drugs approved by the Cancer Drug fund. The NICE TA's guide clinicians regarding second line therapy following Sorafenib and Levatinib, but third line options are limited in that only mutation specific treatments like Selpercatinib can be considered in a limited number of patients. There is limited guidance on second line and subsequent treatments for non-mutation specific RAI-refractory DTC, with currently only best supportive care offered after lenvatinib or sorafenib. Cabozantinib's data from the COSMIC 311 trail offers a promising avenue of treatment for patients whose disease progresses despite second line therapy with existing TKI therapy. Formalising this recommendation through a NICE technology appraisal will give confidence to clinicians to prescribe this peer reviewed intervention with confidence. This offers patients with radio-refractory differentiated thyroid cancer a credible treatment option compared to best supportive care. Recognising Cabozantinib, through a formal NICE appraisal ,as a second line agent, potentially offers this drug to a wider cohort of patients who stand to benefit from reduced tumour burden and reduced long term disability if the disease is effectively controlled.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Cabozantinib will be used in as a second line drug after other TKI's like Sorafenib and Levatinib. Approval of this technology will be bring UK practice in line with updated guidelines from the European Society of Medical Oncology.

• How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	Oncologists are already familiar with TKI drugs, so no additional expertise is needed in prescribing Cabozantinib for eligible patients. The drug will be exclusively prescribed by secondary care by Thyroid oncologists. No additional investments in terms of facilities should be required as TKI drugs are already being prescribed for other cancers now.
13. 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? • Do you expect the technology to increase health-related quality of life more than current care?	• Based on data from the COSMIC 311 trial, Cabozantinib offers benefit in terms of progression free survival with a signal towards improved overall survival. This drug therefore offers a new treatment option for patients with radioiodine-refractory DTC who have no other available standard of care, apart from best supportive care. • Based on current data, Cabozantinib has the potential to offer improved and increased health related quality of life, through improved progression free survival. • Results from the COSMIC 311 trial showed that Cabozantinib significantly reduced the risk of disease progression or death versus placebo with a median progression free survival of 11.0 months compared to 1.9 months with placebo.

14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Treatment with Cabozantinib would benefit patients with radioiodine refractory differentiated thyroid cancer who have progressive metastatic disease. As women have a higher prevalence of differentiated thyroid cancer compared to men, offering this treatment to women with progressive and metastatic disease would improve outcome in women and address the differential morbidity and mortality that women are exposed to, by virtue of the higher prevalence of the disease in women.
15. 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)	Cabozantinib is recommended as a second line agent in the treatment of progressive and radio-resistant DTC by the European Society of Medical Oncology. Using Cabozantinib in the eligible UK population with progressive and radio-resistant DTC, should not pose any added significant burden on prescribers and patients. Patients will obviously need to be closely monitored but the tolerability profile from the COSMIC trial did not suggest that the treatment was poorly tolerated. Using Cabozantinib as a second line agent in progressive DTC offers UK oncologists an additional drug in their armamentarium to treat non-responsive metastatic disease. ECG and blood profile monitoring of patients on treatment will be required as part of standard care but the impact and burden of this on resources and oncologist time is unlikely to be significant. There are likely to be resource implications for cross sectional imaging studies and radiology investigations to assess disease response as additional periodic radiology scans are likely to be required to assess disease response rates as per RECIST criteria.

16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Treatment will be used as a second line agent in metastatic, progressive and radio-resistant DTC in eligible patients. Treatment and disease response should be reviewed periodically to ensure ongoing benefit and patient tolerability. Cross sectional radiology scans will be needed periodically to assess to therapy.
17. 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? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	Substantial health benefits are likely to be realised based on data from the COSMIC trial which demonstrated that Cabozantinib significantly prolonged progression-free survival in patients with radio-resistant and progressive metastatic disease. The administration of the medication is not onerous as the drug can be given orally once daily. I do not have the expertise to comment on QALY metrics.
18. 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? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Current second line treatment options for progressive (and radioresistant) DTC are limited in the UK, with patients getting best supportive care after treatment failures with drugs like Sorafenib and Levatinib. Treatment with Cabozantinib offers a credible therapeutic option to patients in this group and the safety and efficacy of the drug has been evaluated in the COSMIC 311 Trial. Cabozantinib showed significant improvement in progression-free survival over placebo. Cabozantinib therefore is a viable therapeutic option with the potential to reduce disease burden and improve progression free survival. A drug that offers this

	therapeutic response in eligible patients is clearly a significant step compared to offering only best supportive care to patients with progressive metastatic radio-resistant disease.
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	TKI's like Cabozantinib are not without significant adverse effects and treatment will need to be monitored closely. Data from the COSMIC trial showed that Grade 3 or 4 adverse events occurred in 71 (57%) of 125 patients receiving cabozantinib and 16 (26%) of 62 receiving placebo, the most frequent of which were palmar–plantar erythrodysaesthesia (13 [10%] *vs* 0), hypertension (11 [9%] *vs* 2 [3%]), and fatigue (ten [8%] *vs* 0). Serious treatment-related adverse events occurred in 20 (16%) of 125 patients in the cabozantinib group and one (2%) of 62 in the placebo group. There were no treatment-related deaths. Adverse effects reported in the COSMIC 311 Trial were manageable and consistent with the known safety profile of Cabozantinib. Oncologists are familiar with managing the adverse effects of TKI drugs. Dose adjustment with lower doses can mitigate adverse effects but arguably the benefit in terms of progression free survival poses a strong argument for using the drug in eligible patients with close monitoring of adverse effects.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting?	From the clinical standpoint, the most relevant outcome data from COSMIC 311are:

• What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	1. Cabozantinib significantly reduced the risk of disease progression or death versus placebo 2. The median Progression Free Survival was 11.0 months in the cabozantinib treated patients compared to 1.9 months in those receiving placebo. 3. The Objective Response Rate as estimated by RECIST criteria was 18% in the cabozantinib group and 0% in the placebo arms. 4. Adverse effects reported in the COSMIC trail were manageable and consistent with the known safety profile of cabozantinib.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
22. How do data on real-world experience compare with the trial data?	Demographic data from the COSMIC-311 trial in terms of age, sex and ethnic composition is similar to what one would encounter in most Thyroid cancer MDT's as is the proportion of patients who have RAI refractory DTC Adverse effects reported in the COSMIC Trial were consistent with the known safety profile of Cabozantinib.
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil	No specific inequities are immediately apparent. This technology is likely to benefit both sexes but as the prevalence of DTC is higher in women, this group of patients stands to gain from access to a treatment and potentially reduce the differential excess morbidity in women, just by virtue of more women having the disease.

partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Uncertainty around the effect of cabozantinib on overall survival Is it plausible that all people in the "best supportive care" treatment arm who are alive at 5 years will die at this timepoint? Is it plausible that cabozantinib continues to have the same benefit compared to best supportive care for the full duration of the model (35 years)?	Overall survival was not a primary end point in the COSMIC 311 trail and interpretation of survival was confounded by the cross over design of the trail. However, there appeared to be a 'signal' towards better survival compared to placebo. Patients on best supportive care are most likely to die within 3-5 years of disease progression. However it is plausible that patients surviving more than 5 years, may have different tumour biology characteristics. For patients who tolerate treatment it is plausible that cabozantinib continues to have the same benefit compared to best supportive care for the full duration of the model, though in reality patients with radioiodine refractory, metastatic DTC are unlikely to survive that long.

Uncertainty around the most appropriate health state utility values • Do you think it's plausible that people who have pre-progressed disease could have a higher utility value than the UK general population? • Do you think the pre-progression utility values collected in COSMIC-311 are more/less appropriate to use than the utility values which have been used in the majority of previous NICE appraisals of treatments for thyroid cancer (i.e., Fordham et al.)	Unable to comment on this domain as this is outside my expertise
Issues relating to resource use and costs • Should the costs of post-progression cabozantinib be included in the economic model, given	There are arguments for looking at the economic modelling both ways. • Patients may include a cohort whose disease initially progresses and then stabilises on Cabozantinib. Economic modelling should be undertaken in this group to assess the cost impact of long-term therapy • Additional modelling should also assume that treatment costs will stop at the point of disease progression. • The ECG frequency appears reasonable given the known adverse profile of Cabozantinib

that in the marketing authorisation and the COSMIC-311 trial, treatment with cabozantinib could continue post-progression? Or should the model assume that all treatment costs stop at the point of progression? • Should wastage costs be included in the model, given that people who stop treatment due to progression or death before completing a full pack of treatment will incur some level of wastage. • The company's model assumes that people taking cabozantinib will undergo an electrocardiogram (ECG) once every 6 months – is this accurate or would they be more frequent?	• I would suggest that costs relating to cross sectional imaging CT/MRI scans be factored in as patients will need additional scans to assess stability(or not) of disease progression. • Not unreasonable to include wastage costs as the drug is expensive,

Are there any important issues that have been missed in EAR?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Progressive, metastatic and radioresistant differentiated thyroid cancer is associated with significant morbidity and reduced life expectancy.

Results from the COSMIC 311 trial showed that Cabozantinib significantly reduced the risk of disease progression or death versus placebo with a median progression free survival of 11.0 months compared to 1.9 months with placebo.

Current therapeutic options for treating RAI-refractory DTC in the UK are limited with best supportive care the only option after using other TKI's like Sorafenib and Levantinib

Data from the COSMIC 311 trail positions Cabozantinib as a potential new treatment option for previously treated RAI-refractory DTC with the potential to reduce disease burden and reduce disease progression in non-mutation specific radio-iodine refractory differentiated thyroid cancer. This therefore offers hope to patients whose only other option would have been best supportive care after second line TKI therapy.

Thank you for your time.

NICE National Institute for Health and Care Excellence

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☒ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Clinical expert statement and technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing 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. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

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. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **Wednesday 25th January 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Treating thyroid cancer and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Professor Jonathan Wadsley
2. Name of organisation	NCRI Thyroid Cancer Group
3. Job title or position	Consultant Clinical 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 thyroid cancer? ☐ A specialist in the clinical evidence base for thyroid cancer 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 did not submit one, I do not 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
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None
8. What is the main aim of treatment for differentiated thyroid cancer which is unsuitable for or refractory to	To stop progression to help maintain quality of life and to extend survival

radioactive iodine and has progressed during or after prior systemic therapy? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	
9. 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)	Disease being held stable (but not necessarily shrinking) can be a clinically significant outcome.
10. In your view, is there an unmet need for patients and healthcare professionals in differentiated thyroid cancer which is unsuitable for or refractory to radioactive iodine and has progressed during or after prior systemic therapy?	Yes. Except for the very small subset of patients who have targetable genetic alterations (NTRK and RET fusions) there is no other active treatment available, and prognosis is very poor.
11. How is differentiated thyroid cancer which is unsuitable for or refractory to radioactive iodine and has progressed during or after prior systemic therapy currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • 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.) • What impact would the technology have on the current pathway of care?	Except for the small group alluded to above with specific targetable genetic alterations, these patients are treated with best supportive care, which may include palliative radiotherapy and specialist palliative care input. There are no specific clinical guidelines- this would be considered part of standard oncological care. There are no significant difference in opinion regarding this. The availability of cabozantinib would open another line of active treatment for this group of patients, extending progression free survival, and potentially overall survival.

12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	This will be an additional treatment option, so would require additional healthcare resource use over current care. The treatment would only be used in specialist thyroid oncology clinics, usually a tertiary care arrangement. Thyroid oncologists are already familiar with using this drug for another licensed indication and numbers of patients are very small so further training, facilities or equipment are unlikely to be required.
13. 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? • Do you expect the technology to increase health-related quality of life more than current care?	Data regarding the effect of this treatment on overall survival are difficult to interpret for all of the reasons raised in the EAG report, although I think there is likely to be a benefit for this group who otherwise have an extremely poor prognosis. There is however firm evidence that the treatment extends progression free survival without impairing HRQOL so yes, I do expect it to increase HRQOL.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	To my knowledge, no subgroups of patients have been identified who are more or less likely to benefit.

15. 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)	Patients will require specialist monitoring in a thyroid oncology clinic whilst on treatment. This will require at least 4 weekly clinic review with blood tests, 12 weekly CT scans, access to out of hours oncology support in case of development of toxicity
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Patients are likely to start treatment on evidence of progression on a previous line of treatment. They will stop treatment when they are no longer felt to be deriving clinical benefit. Neither of these should result in additional testing.
17. 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? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	No
18. 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? • Is the technology a 'step-change' in the management of the condition?	Yes- currently there is no active treatment available for the majority of these patients and their prognosis is very poor, so it is addressing a significant unmet need.

• Does the use of the technology address any particular unmet need of the patient population?	
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Patients require careful monitoring in a specialist thyroid cancer oncology clinic to ensure that side effects are adequately managed, either with supportive medications or dose reduction. It is usually possible to reach a dose level at which patients are deriving benefit without troublesome toxicity.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Yes- the COSMIC-311 trial was conducted in exactly the setting in which we would plan to use this treatment in the UK, and if approved the treatment would be used in this setting. An improvement in progression free survival is an important outcome for this group of patients since if disease is not progressing they are unlikely to develop new disease-related symptoms. I am not aware of any adverse effects not previously reported.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
22. How do data on real-world experience compare with the trial data?	The only experience I have of using this treatment is with patients who were taking part in the COSMIC-311 trial.
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of	No.

people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	I note comments regarding thyroid cancer being more common in women. However, as noted in the EAG report, men with thyroid cancer tend to have a worse prognosis and therefore the numbers of men and women with the type of aggressive disease that requires this treatment are approximately equal- as demonstrated in the trial population.

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Uncertainty around the effect of cabozantinib on overall survival Is it plausible that all people in the "best supportive care" treatment arm who are alive at 5 years will die at this timepoint? Is it plausible that cabozantinib continues to have the same benefit compared to best supportive care for the full duration of the model (35 years)?	A vertical drop in survival at 5 years, as suggested by the company's model, is not plausible. However, in my view, neither is it plausible that 10% of patients who have progressed on Lenvatinib and not received further treatment will still be alive at 5 years. Whilst it is true that metastatic differentiated thyroid cancer has very variable rates of progression, the patients being considered for this treatment are a very select subgroup, by definition with a poorer prognosis. They will already have had evidence of significant disease progression to justify first line therapy with Lenvatinib, and will then have progressed through that treatment in order to be considered for second line treatment. In my personal experience I cannot think of a single patient who has survived more than 2 years beyond progression on Lenvatinib, unless further therapy has been available. The fundamental problem is of inadequate data to model overall survival, especially in the BSC arm of the COSMIC trial. I note that of the 62 patients in the placebo arm in the published analysis, there were no patients actually at risk at the 1 year time point (accepting that 48 were censored). Given this

	very limited number of patients and very short follow up for the majority, I think it is unrealistic to expect to be able to model overall survival in any meaningful way, and perhaps not surprising that attempts to model the longer-term overall survival lack any credibility.
Uncertainty around the most appropriate health state utility values • Do you think it's plausible that people who have pre-progressed disease could have a higher utility value than the UK general population? • Do you think the pre-progression utility values collected in COSMIC-311 are more/less appropriate to use than the utility values which have been used in the majority of previous NICE appraisals of treatments for thyroid cancer (i.e., Fordham et al.)	I agree with the EAG that it is not plausible that this group of patients could have a higher utility value than the UK general population. The patient group recruited to COSMIC-311 is different to that considered in other NICE appraisals of thyroid cancer treatments, in that all patients in this study were receiving second line treatment, whereas in prior studies many patients were receiving first line treatment and therefore may have had a lower symptom burden. Therefore I think it would be more appropriate to use the utility values collected in this study.

Issues relating to resource use and costs • Should the costs of post-progression cabozantinib be included in the economic model, given that in the marketing authorisation and the COSMIC-311 trial, treatment with cabozantinib could continue post-progression? Or should the model assume that all treatment costs stop at the point of progression? • Should wastage costs be included in the model, given that people who stop treatment due to progression or death before completing a full pack of treatment will incur some level of wastage. • The company's model assumes that people	I agree that the costs of post-progression cabozantinib should be included in the economic model, as the reality is that in the absence of any other lines of treatment, it is likely that patients would continue on cabozantinib for as long as they are considered to be deriving clinical benefit, and this may extend beyond first evidence of radiological progression. I agree that wastage costs should be included as this loss is almost inevitable, and the potential cost not insignificant. Regarding frequency of ECGs, the SmPC for cabozantinib does not specify a desirable frequency. I suspect there is significant variability of practice across the UK here. In my personal practice I would typically undertake an ECG at baseline, at 1 month, and assuming no changes of concern then probably only once every 6 months.

taking cabozantinib will undergo an electrocardiogram (ECG) once every 6 months – is this accurate or would they be more frequent?	
1	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

The evidence suggests that cabozantinb can provide meaningful clinical benefit to patients with advanced iodine refractory differentiated thyroid cancer who have progressed on a prior tyrosine kinase inhibitor by extending progression free survival.
Due to the immaturity of the data available it is very difficult to comment on the magnitude of the likely survival benefit from this treatment.
It is therefore very challenging to comment on the exact cost effectiveness of this treatment at this stage.
Click or tap here to enter text.
Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **Wednesday 25th January 2023**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we

received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	NCRI-ACP-RCP-RCR
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Table 1 About you

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Uncertainty around the effect of cabozantinib on overall survival	No	Our experts believe that a vertical drop in survival at 5 years, as suggested by the company's model, is not plausible. However, neither is it plausible that 10% of patients who have progressed on Lenvatinib and not received further treatment will still be alive at 5 years. Whilst it is true that metastatic differentiated thyroid cancer has very variable rates of progression, the patients being considered for this treatment are a very select subgroup, with a poorer prognosis. They will already have had evidence of significant disease progression to justify first line therapy with Lenvatinib and will then have progressed through that treatment to be considered for second line treatment. Our experts have not experienced any examples of patients surviving more than 2 years beyond progression on Lenvatinib unless further therapy has been available. The fundamental problem is of inadequate data to model overall survival, especially in the BSC arm of the COSMIC trial. Our experts note that of the 62 patients in the placebo arm in the published analysis, there were no

		patients at risk at the 1-year time point (accepting that 48 were censored). Given this very limited number of patients and very short follow up for the majority, our experts believe it is unrealistic to expect to be able to model overall survival in any meaningful way, and perhaps not surprising that attempts to model the longer-term overall survival lack any credibility.
Uncertainty around the most appropriate health state utility values	No	Our experts agree with the EAG that it is not plausible that this group of patients could have a higher utility value than the UK general population. The patient group recruited to COSMIC-311 is different to that considered in other NICE appraisals of thyroid cancer treatments, in that all patients in this study were receiving second line treatment, whereas in prior studies many patients were receiving first line treatment and therefore may have had a lower symptom burden. Therefore, our experts believe it would be more appropriate to use the utility values collected in this study.
Issues relating to resource use and costs	No	Our experts agree that the costs of post-progression cabozantinib should be included in the economic model, as the reality is that in the absence of any other lines of treatment, it is likely that patients would continue cabozantinib for as long as they are considered to be deriving clinical benefit, and this may extend beyond first evidence of radiological progression. Our experts agree that wastage costs should be included as this loss is almost inevitable, and the potential cost not insignificant. Regarding frequency of ECGs, the SmPC for cabozantinib does not specify a desirable frequency. Our experts suspect there is significant variability of practice across the UK. One of our experts noted that in their personal practice they would typically undertake an ECG at baseline, at 1 month, and assuming no changes of concern then probably only once every 6 months.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]

Addendum: EAG comments on the company's technical engagement response

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK Kate (Shijie) Ren, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Aline Navega Biz, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Sarah (Sa) Ren, Research Associate, ScHARR, University of Sheffield, Sheffield, UK
Correspondence Author	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK
Date completed	6th February 2023

1. Introduction

1.1 Summary of company's TE response

This addendum provides a summary and critique of the company's technical engagement (TE) response[1] for the appraisal of cabozantinib for previously treated differentiated thyroid cancer (DTC) unsuitable for or refractory to radioactive iodine (RAI). Both the company's TE response and this addendum should be read alongside the company's submission (CS)[2] and the External Assessment Group (EAG) report.[3]

The company's TE response consists of a written document[1] and a revised version of the company's executable model, which includes most of the amendments applied in the EAG's preferred analysis. The company's written response includes discussion around the three key issues raised in the EAG report.[3] In addition, the company has amended the model population for the base case analysis to reflect the second-line (2L) subgroup of the COSMIC-311 trial.[4] This additional analysis is presented as part of the company's response to Issue 1; however, this change in modelled population is considered as a separate issue within this EAG addendum. The key issues, the additional evidence presented in the company's TE response and the company's additional amendments to the economic model are summarised in Table 1.

The additional evidence, analysis and discussion relating to these issues are summarised and critiqued in Section 2 of this addendum. Additional economic analyses undertaken by the EAG are presented in Section 3.

Table 1: Summary of key issues, additional evidence and updates to the company's model

ID4046	Summary of issue	Is additional evidence presented?	Has the EAG's preferred model been amended?
Additional issue	DTC population included in model	PFS, OS and TTD data presented for 2L subgroup in COSMIC-311	Yes – model amended to use data from 2L subgroup in COSMIC-311(see Table 2).
Issue 1	Uncertainty around the effect of cabozantinib on overall survival	Additional clinical input on expected survival for BSC and cabozantinib	Yes – model amended to use data from 2L subgroup in COSMIC-311 (see Table 2). Additional blended survival analysis applied for BSC group only in scenario analysis.
Issue 2	Uncertainty around the most appropriate health state utility values	None	Yes – updated model uses adjusted utility values† from Fordham *et al.*[5] applied (see Table 2). Utility cap removed.
Issue 3	Issues relating to resource use and costs	None	Yes - RDI applied instead of compliance. TTD modelled using generalised gamma function (see Table 2).

EAG - External Assessment Group; DTC - differentiated thyroid cancer; PFS - progression-free survival; OS - overall survival; TTD - time to treatment discontinuation; BSC - best supportive care; 2L - second-line; RDI - relative dose intensity
† Adjusted for educational qualification level and EQ-5D-3L (usual activities and anxiety/depression) ratings using UK norms.

1.2 Description of company's TE model amendments

Table 2 summarises the differences between the company's original base case model,[2] the EAG's preferred model[3] and the company's updated TE base case model.[1] The differences between the company's TE model and the EAG's preferred model are:

- The company's TE model reflects the 2L subgroup of COSMIC-311.[4] All parametric survival models have been re-fitted to the data for this subgroup. All previous survival analyses used in the company's original model and the EAG's preferred analyses reflected the intention-to-treat (ITT) population.
- The company's TE model uses a generalised gamma distribution to model time to treatment discontinuation (TTD) in the 2L subgroup. The EAG's preferred model in the ITT population used a Weibull distribution.
- The company's updated model uses adjusted utility values from Fordham *et al.*[5] and does not cap health state utility values by general population utility. The EAG's preferred model uses unadjusted utility values from Fordham *et al.* and includes a general population utility cap.
- The company's TE model adjusts drug costs according to relative dose intensity (RDI) in COSMIC-311. The EAG's preferred model adjusted drug costs based on compliance to account for the flat pricing structure of cabozantinib across different dosage strengths.

Table 2: Summary of company's original and updated base case models and EAG preferred model

Aspect of model	Company's original model[2]	EAG's preferred model[3]	Company's TE base case model[1]
Population	COSMIC-311 ITT population	COSMIC-311 ITT population	COSMIC-311 2L subgroup
OS	Exponential. BSC 5-year death assumption included.	Exponential. BSC 5-year death assumption excluded.	Exponential. BSC 5-year death assumption excluded.
PFS	Weibull	Weibull	Weibull
TTD	Exponential. Capped by PFS.	Weibull. PFS cap removed.	Generalised gamma. PFS cap removed.
Health state utility values	Fordham *et al.*[5] Adjusted values† (PF=0.87, PD=0.52)	Fordham *et al.*[5] Unadjusted values (PF=0.80, PD=0.50)	Fordham *et al.*[5] Adjusted values† (PF=0.87, PD=0.52)
General population utility cap	Not included	Included	Not included
Drug costing approach	Compliance	Compliance	RDI
Wastage	Excluded	Included	Included

EAG - External Assessment Group; TE - technical engagement; ITT - intention-to-treat; 2L - second-line; OS - overall survival; PFS - progression-free survival; TTD - time to treatment discontinuation; PF - progression-free; PD - progressed disease; RDI - relative dose intensity; BSC - best supportive care

† Adjusted for educational qualification level and EQ-5D-3L (usual activities and anxiety/depression) ratings using UK norms

1.3 Summary of company's TE base case results

Unless otherwise stated, all incremental cost-effectiveness ratios (ICERs) presented in this EAG addendum exclude quality-adjusted life year (QALY) weighting. ICERs including QALY weighting can be found in Appendix 1.

The company's updated base case results for the 2L subgroup are summarised in Table 3. The deterministic version of the company's model suggests that the ICER for cabozantinib versus best supportive care (BSC) is £23,050 per QALY gained. The probabilistic ICER is higher at £25,081 per QALY gained. The company's TE response also reports the results of a range of scenario analyses for the 2L subgroup. The ICERs estimated from the deterministic scenario analyses range from £18,451 to £30,218 per QALY gained. The ICERs estimated from the probabilistic scenario analyses (including QALY weighting) range from £10,122 to £23,552 per QALY gained; the results of the probabilistic scenario analyses excluding QALY weighting are not provided in the company's TE response and have not been re-run by the EAG.

Table 3: Company's updated base case analysis, 2L subgroup (excluding EAG's error correction)

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER	DM
Probabilistic model†								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£25,081**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
Deterministic model								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£23,050**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	

2L - second-line; LYG - life year gained, QALY - quality-adjusted life year, ICER - incremental cost-effectiveness ratio; DM - decision modifier; BSC - best supportive care

** Undiscounted*

† Results generated by the EAG by re-running the company's PSA sub-routine

2. Summary of company's TE response and EAG critique

2.1 Additional issue: Change in modelled population

The company has amended the population reflected in the survival analysis and economic model to focus specifically on the 2L subgroup of the COSMIC-311 trial.[4] The company's TE response[1] (page 6) states that this analysis has been performed *"to alleviate the EAGs concern with application of a 5-year death assumption."* The EAG is unclear how amending the model population addresses uncertainty around the model predictions for the BSC group, but notes that this change has potential implications for the relative clinical effectiveness and cost-effectiveness of cabozantinib for the treatment of RAI-refractory DTC.

The company's TE response[1] reports progression-free survival (PFS), overall survival (OS) and time to treatment discontinuation (TTD) outcomes for the 2L population; the hazard ratios (HRs) for PFS and

OS for the ITT population and the 2L subgroup are presented in Table 4. The company has also updated all of the survival analyses and treatment switching adjustment analyses using the data for the 2L subgroup. The parametric survival distributions for PFS and OS selected by the company for the 2L subgroup are the same as those used in the company's original model[2] (Weibull and exponential models, respectively; see Table 2), whilst the TTD function has been changed from the Weibull to the generalised gamma distribution. Comparisons of the Kaplan-Meier estimates and the model predictions for PFS, OS and TTD in the 2L subgroup are reported in Figures 1-5 of the company's TE response. Goodness-of-fit statistics (the Akaike Information Criterion [AIC] and Bayesian Information Criterion [BIC]) for each model are provided in Tables 4-6 of the company's TE response. Hazard plots and log cumulative hazard plots are not presented. For brevity, the company's survival plots and AIC/BIC statistics are not reproduced here.

Table 4: Summary of stratified HRs for full ITT population and 2L subgroup in COSMIC-311, CCO2

Model population	PFS	OS (RPSFT-adjusted)*
Full ITT population (N=258)	[redacted]	0.65 (0.28, 1.53)
2L subgroup (N=191)	[redacted]	[redacted]

ITT - intention-to-treat; 2L - second-line; PFS - progression-free survival; OS - overall survival; RPSFTM - Rank-Preserving Structural Failure Time; N - number
** The original analysis for the ITT population reports the 95% CI, whilst the TE response reports the 96% CI*

EAG comments

With respect to the company's decision to re-focus the analysis on the 2L subgroup, the EAG notes the following:

- The 2L subgroup of COSMIC-311[4] is likely to better reflect the population of RAI-refractory DTC patients who would receive cabozantinib in the NHS in England. As this is a subgroup, the sample size is reduced, leading to greater uncertainty in the model predictions.
- The point estimates of the HRs for PFS and OS between cabozantinib and BSC in the 2L subgroup appear to be better than those estimated for the full ITT population (see Table 4).
- Whilst the company's TE response indicates that the company's decision to focus on the 2L subgroup has been undertaken in an attempt to alleviate the EAG's concerns regarding the 5-year death assumption used in the BSC group of the original model, this change has virtually no bearing on the OS extrapolation for the BSC group. The EAG has plotted the model predictions for PFS and OS in the ITT population and the 2L subgroup in Figure 1 and Figure 2, respectively. As shown in these plots, the predictions for BSC are almost identical regardless of the population considered – the company's original model for the ITT population suggested a 5-year OS probability of approximately [redacted], whilst the company's TE model for the 2L subgroup also suggests a 5-year OS probability of approximately [redacted].

- Within COSMIC-311, randomisation was stratified by previous lenvatinib (yes vs. no) and age (≤65 years vs. >65 years). The number of prior lines of therapy was not included as a stratification factor. The company's TE response does not provide any information on baseline characteristics in the 2L subgroup. It is unclear whether the treatment groups are well balanced within the subgroup.
- Following receipt of the company's TE response, the EAG asked the company to clarify whether they are now seeking a positive recommendation for cabozantinib specifically for patients at 2L. In their response, the company stated that ideally, they would like to receive a positive recommendation for the full licensed population, including patients at 2L and subsequent lines. The company's response also comments that cabozantinib demonstrates improved cost-effectiveness in the 2L population. Whilst it is the company's intention to obtain a positive recommendation for the full licensed indication, the updated TE model now reflects a narrower subgroup.
- The EAG was unable to fully critique the updated survival analyses for the 2L subgroup as the company's TE response does not provide hazard plots or log-cumulative hazard plots. The company's model selections for PFS and OS remain unchanged from original base case in the ITT population.
- The company has amended the parametric survival model for TTD - the company originally selected the exponential model including a PFS cap, whereas the EAG's preferred analysis used the Weibull model without the PFS cap. The company's TE response states that the Gompertz and generalised gamma models were closest to the PFS curve and that this aligns with the wording of the Summary of Product Characteristics (SmPC) for cabozantinib:[6] *"patients should continue treatment until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs."* The TE response (page 17) also states that "*Based on the goodness of fit and visual inspection against observed KM data, the generalised gamma distribution has the lowest AIC and BIC and best fit to KM data compared with Gompertz."* The EAG does not believe that the wording of the SmPC, which permits post-progression treatment with cabozantinib, is a strong rationale for assuming that TTD must be similar to PFS. The EAG also notes that the generalised gamma model provides a notably worse fit than several other survival models (including the exponential, Weibull and Gompertz distributions) when judged according to BIC. As shown in Figure 3, the generalised gamma TTD function is consistently below the PFS function at all timepoints – this implies that no patient receives post-progression cabozantinib. The EAG believes that the exponential or Weibull models may be more appropriate than the generalised gamma model, but further clinical input would be helpful to determine the plausibility of these functions.

Figure 1: Comparison of model-predicted PFS for 2L subgroup and ITT population, CCO2 (generated by the EAG)

2L - second-line ; ITT - intention-to-treat; CCO - clinical cut-off; PFS - progression-free survival; Cabo - cabozantinib; BSC - best supportive care

Figure 2: Comparison of model-predicted OS for 2L subgroup and ITT population, CCO2 (generated by the EAG)

2L - second-line ; ITT - intention-to-treat; CCO - clinical cut-off; OS - overall survival; Cabo - cabozantinib; BSC - best supportive care

Figure 3: Comparison of modelled TTD and PFS for 2L subgroup, CCO2 (generated by the EAG)

TTD - time to treatment discontinuation; PFS - progression-free survival; 2L - second-line; CCO - clinical cut-off; TE - technical engagement

2.2 Issue 1: Uncertainty around the effect of cabozantinib on overall survival

The company's TE response[1] presents additional expert opinion on expectations of OS in patients with RAI-refractory DTC receiving cabozantinib or BSC at 2, 5 and 10 years. These estimates have been used to inform parametric survival model selection for PFS and OS in the company's updated 2L base case model. The company has also undertaken a blended survival analysis which uses the mean of the elicited values for BSC across all experts; this is presented as a scenario analysis.

Summary of clinical experts' estimates of OS for cabozantinib and BSC

Table 5 summarises the clinical experts' estimates of OS at 2, 5 and 10 years obtained by the EAG and the company. The table also includes the model predictions obtained from various iterations of the company's economic model.

Table 5: Clinical expert estimates of OS for cabozantinib and BSC and company's model predictions (adapted from company's TE response, Table 1)

	Cabozantinib			BSC		
	2 years	**5 years**	**10 years**	**2 years**	**5 years**	**10 years**
Clinical experts' estimates						
EAG Advisor 1	63%	35%	0%	50%	10%	0%
EAG Advisor 2	45%	25%	0%	35%	13%	0%
Company's clinical advisors[2]	-	-	-	-	[redacted]	-
Company Advisor 1*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company Advisor 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company Advisor 3	-	-	-	[redacted]	[redacted]	[redacted]
Company Advisor 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean of all estimates	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Model predictions‡						
Company's TE base case (2L subgroup[1])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company's model (including BSC 5-year death assumption) (company's original model[2])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company's model (excluding BSC 5-year death assumption) (EAG preferred model[3])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company's model (using CCO2 for cabozantinib and CCO1 for BSC) (company's post-clarification model[7])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Company's model with blended survival scenario analysis (TE sensitivity analysis[1])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BSC - best supportive care; CCO - clinical cut-off; EAG - External Assessment Group; TE - technical engagement
**Mean calculated from two interviews conducted with same clinical advisor*
† The proportion of patients predicted to be alive at 2 years on BSC in the company model excluding BSC 5-year death assumption is [redacted] rather than [redacted] reported in the EAG report.
‡ Some of the OS probabilities presented in Table 1 of the company's TE response differ slightly from those estimated from the EAG's amended trace in the company's model.

Summary of company's blended survival analysis

The company conducted a blended survival analysis incorporating expert clinical opinion and data from the ITT population of COSMIC-311[4] in extrapolation. The analysis was undertaken using methods reported by Che *et al.*[8] The blended survival analysis was conducted only for OS in the BSC arm, including adjustment for treatment switching using the Rank-Preserving Structural Failure Time (RPFST) method.

The blended survival analysis approach consists of two separate processes to extrapolate long-term survival. The first component is similar to the standard survival extrapolation whereby a parametric model is used to fit to the observed data with the main objective being to produce the best fit to the observed data. The second component is the blending process. A separate external survival curve is produced based on the external data. The blended survival curve is obtained as a function of both components weighted using the cumulative distribution function of a beta random variable. This beta

random variable controls the blending rate. The blended survival curve consists of three intervals: (i) the follow-up interval where the survival curve is identical to the model fitted to the observed data, (ii) the blending interval where the survival curve is a combination of the model fitted to the observed data and the model fitted to the external data, (iii) the long-term interval where the survival curve is identical to the model fitted to the external data.

A piecewise exponential model with time intervals of 2 months was used for RPFST-adjusted OS from the BSC arm. A Weibull model was fitted to the survival estimates provided by the experts at 2, 5 and 10 years (the mean estimates across all of the EAG's and company's advisors) assuming N=70 patients in the base case and N=70±20% in scenario analyses. The base case analysis assumes that the blending interval is from 14 months to 24 months with a scenario analysis assuming that it is from 14 months to 60 months. The base case uses Beta(3,3) for the weight function to estimate the blending rate (assuming a steady monotonic increasing hazard within the blending interval) and the scenario analysis uses Beta(2,5) (assuming a faster monotonic increasing hazard within the blending interval).

The base case for the blended survival analysis produced mean survival estimates of [redacted], [redacted] and [redacted], at 2, 5, and 10 years respectively (see Figure 4). The scenario analysis changing the blending interval and blending rate provided identical mean survival estimates at 2, 5 and 10 years. The scenario analysis assuming a different sample size for the external evidence had only a small impact on the estimates (see company's TE response, Figures 9-12).

Figure 4: Company's blended survival analysis scenario analyses, ITT population, CCO2 (reproduced from company's TE response, Figure 8)

BSC - best supportive care; RPSFT - Rank-Preserving Structural Failure Time; KM - Kaplan-Meier

EAG comments

With respect to the uncertainty around the relative benefits of cabozantinib on OS and the expected OS for BSC, the EAG notes the following:

- The additional expert opinion obtained by the company is useful for informing expected OS in the model. However, no information is provided regarding what the company asked the experts.

This might be important given the company's decision to re-focus the model on the 2L subgroup.

- The estimates provided by the experts indicate wide variation in expectations of OS for both groups. For example, in the cabozantinib group, expected 5-year OS ranges from 8% to 35%, whereas in the BSC group, expected 5-year OS ranges from 0% to 13%. Given the variability in responses, taking a mean estimate may not be meaningful.
- The EAG acknowledges that the blended survival analysis approach has the potential to address the issues in the extrapolation of OS using standard survival models. However, the company's blended survival curve does not fit the observed data well and also overestimates the mean survival estimates provided by the experts (see Table 5). The EAG suspects that there are alternative survival models for both the observed data and experts' estimates which could be more suitable and which would lead to an improvement in the overall fit. The EAG also questions the company's choice to apply the blended survival analysis to the BSC arm but not to the cabozantinib arm, as the current fit to the observed data in the tail area of the cabozantinib arm is poor (see EAG report,[3] Figure 17). The EAG is also unsure why an equivalent blended analysis has not been presented for the 2L subgroup.
- The EAG has estimated the time-varying HR from the blended survival analysis using the cumulative survival probabilities contained in the company's economic model (see **Figure 5**). As shown in the plot, the HR appears to generally decrease (improve) over time. This contrasts with the HR estimated for the ITT population which appeared to be worsening over time (see EAG report, Figure 24). The EAG believes that the company's blended survival analysis may be optimistic.

Figure 5: Estimated time-varying HR in company's blended survival analysis scenario analysis, ITT population, CCO2 (generated by the EAG)

HR - hazard ratio
Note: Plot generated by approximating the hazard functions for the BSC group (based on the average of the sampled survivor functions obtained from the blended survival analysis) and the cabozantinib group (the exponential model from the trace)

The EAG also notes that the company's TE response[1] states *"the company maintain that the inclusion of the 5-year BSC death assumption is appropriate compared to excluding this assumption."* However, the 5-year death assumption has been excluded from the company's updated TE base case (see Table 2). The EAG is unsure whether this is intentional or an error on the part of the company.

2.3 Issue 2: Uncertainty around the most appropriate health state utility values

The company's TE response[1] argues that the adjusted utility values reported by Fordham *et al.*[5] are the most appropriate values to use in the base case analysis. The company's response also states that the utility values obtained from COSMIC-311[4] lack validity. In addition, the company agrees that an age-adjusted general population utility cap should be applied to ensure that health-related quality of life (HRQoL) in the RAI-refractory DTC population cannot exceed HRQoL in the general population.

EAG comments

The EAG has the following concerns regarding the utility values applied in the company's TE model:

- As described in Section 5.5 of the EAG report,[3] the EAG's preferred model included a constraint which prevented the utility value for the progression-free health state from exceeding the estimated Euroqol 5-Dimensions 5-Level (EQ-5D-3L) utility value for the age- and sex-matched general population.[9] This constraint has been overwritten in the company's TE model (see executable model, worksheet "Quality Of Life Inputs", cells D12:D13). Given the company's preference for using the adjusted values from Fordham *et al.,*[5] this means that the company's TE model suggests that RAI-refractory DTC patients who are progression-free have a higher level of HRQoL compared with the general population (progression-free utility = 0.87 versus general population utility = 0.82). The EAG believes that this is an error which was resolved in the EAG's preferred model, but which has been reintroduced in the company's TE model.
- The EAG believes that it may be reasonable to use the observed utility values reported by Fordham *et al.*[5] However, given that EQ-5D data are available from COSMIC-311,[4] it is also reasonable to consider these values, at least in sensitivity analyses. The EAG notes that the TE responses from the clinical stakeholders suggest that it would be more appropriate to use the estimates from COSMIC-311 rather than Fordham *et al.*[5]
- Three previous NICE appraisals of treatments for thyroid cancer have applied utility values from Fordham *et al.*[5] (TA516, TA550 and TA742[10-12]). Each of these three appraisals used the unadjusted utility values (progression-free utility = 0.80, post-progression utility = 0.50). The EAG believes that using the higher adjusted utility values from this source would create an inconsistency with earlier TAs with no clear rationale for doing so.

2.4 Issue 3 - Issues relating to resource use and costs

The company's TE response[1] provides further discussion around six issues relating to resource use and costs: (i) the inclusion of post-progression costs for cabozantinib; (ii) the inclusion of drug cost adjustments using RDI rather than compliance; (iii) the inclusion of drug wastage costs; (iv) the frequency of electrocardiograms (ECGs); (v) the frequency of computerised tomography (CT) scans and (vi) the costs of concomitant medications. These six issues are discussed and critiqued below.

(i) Inclusion of post-progression costs

The company's TE response[1] states that the company agrees that the costs of post-progression cabozantinib should be included in the economic analysis and that this reflects the intention for cabozantinib to be used in line with its licence. As noted in Section 2.1, the company has selected the generalised gamma to model TTD.

The generalised gamma model for TTD remains lower than the modelled PFS function at all timepoints. This implies that patients do not receive post-progression cabozantinib. However, at CCO2 in COSMIC-311,[4] 6.5% of patients randomised to cabozantinib had received post-progression cabozantinib. The company's selected TTD model therefore appears inconsistent with the COSMIC-311 trial. The EAG also notes that the clinical stakeholders at TE commented that it is likely that in practice, some patients will continue on cabozantinib beyond radiological progression if they are still deriving clinical benefit from treatment.

(ii) Inclusion of drug cost adjustments using RDI

The company's TE response[1] argues that drug costs should be adjusted using RDI rather than compliance. Specifically, the company argues that: (a) the compliance estimate was based on CCO1 whereas an RDI estimate is available from CCO2 and (b) previous NICE appraisals have included the adjustment of drug acquisition costs using RDI regardless of whether the technology has a flat pricing structure across different dosage strengths. The company argues that consistency should be maintained with NICE/EAG standards.

The EAG believes that given the flat pricing structure for cabozantinib, it is more appropriate to adjust cabozantinib costs according to the proportion of days on which patients received treatment (compliance), rather than the average amount of the planned dose received (RDI). For example, if a patient received treatment with 60mg cabozantinib for 15 days followed by treatment with a lower dose of 20mg cabozantinib for 15 days, there would be no cost savings to the NHS. The compliance-adjusted cost would appropriately reflect this scenario, whereas adjusting drug costs by RDI would erroneously suggest a cost saving of 33%. Regardless of precedents in previous NICE appraisals, the EAG believes that using compliance is more appropriate in this case. However, the EAG agrees that it is not ideal for

the compliance estimate to be taken from CCO1 rather than CCO2. The EAG was unable to find the exact compliance estimate of [redacted] applied in the company's model in the Clinical Study Report (CSR) for CCO1, but notes that non-compliance and dosing errors are reported in both CSRs for CCO1 and CCO2[13, 14] and both reports suggest similar numbers of patients who had any dose interruption. The EAG believes that the original compliance estimate from CCO1 might be reasonable to use, but that this could be improved by the company recalculating compliance at CCO2.

(iii) Drug wastage costs

The company agrees with the EAG's approach to modelling wastage (costs incurred once per patient rather than costs incurred in every cycle).

The EAG believes that this issue can be considered resolved.

(iv) Frequency of ECGs

The company argues that ECGs should be assumed to be given every 6 months, as per their base case analysis. The company considers the EAG's additional sensitivity analysis which assumes ECGs are given every 3 months to be inappropriate.

The EAG's analysis around the frequency of ECGs was undertaken in response to comments received from clinical experts and was presented as an additional sensitivity analysis; it does not form part of the EAG's preferred model. The EAG's analyses indicate that the ICER is not sensitive to this parameter.

(v) Frequency of CT scans for BSC

The company argues that the costs of CT scans should be included for all patients receiving BSC. The company considers the EAG's additional sensitivity analysis which excludes the costs of CT scans for patients receiving BSC alone to be inappropriate.

The EAG's analysis around the use of CT scans was undertaken in response to comments received from clinical experts and was presented as an additional sensitivity analysis; it does not form part of the EAG's preferred model. The EAG's analyses indicate that the ICER is not sensitive to this parameter.

(vi) Costs of concomitant medications

The company has considered the inclusion of concomitant medications in the economic model. The company's TE response[1] highlights several limitations associated with the available data on concomitant medications in COSMIC-311[4] and the updated economic model does not include additional costs associated with these treatments. The TE response highlights that similar proportions

of patients in each arm of COSMIC-311 received at least one concomitant medication, that these therapies are all available at very low cost and that the expected impact on the ICER is minimal.

The EAG agrees that including the costs of concomitant medications would be expected to have a minimal impact on the ICER. However, the EAG would have preferred that the company include these costs, particularly as the model suggests that cabozantinib extends OS.

3. Additional analyses undertaken by the EAG

3.1 Verification of the company's TE model

The EAG was able to use the company's TE model to generate the EAG's preferred ICER for the ITT population. However, as noted in Section 2.3, the EAG identified an error whereby the general population utility cap, which was included in the EAG's preferred model, had been erroneously overwritten in the company's TE model. This error is corrected in all exploratory analyses undertaken by the EAG in Section 3.2.

3.2 Additional EAG exploratory analyses using the TE model (excluding QALY weighting)

The EAG undertook additional exploratory analyses using the company's TE model. Despite the company's decision to focus on the 2L subgroup, the EAG's concerns regarding uncertainty around modelled OS remain unchanged (see EAG report, Section 5.3.5, critical appraisal point [4]). For this reason, the EAG's exploratory analyses using the TE model are the same as those presented in the EAG report,[3] but are based on the 2L subgroup rather than the ITT population. Based on the probabilistic version of the model, the EAG's preferred ICER for the 2L subgroup is estimated to be £31,015; the deterministic ICER is similar at £30,218 per QALY gained. As with the company's original model, the ICER remains particularly sensitive to alternative assumptions regarding OS.

The EAG's retains its view that the long-term effect of cabozantinib on OS is highly uncertain and that none of the economic analyses presented by the company or the EAG are ideal. Longer-term follow-up in COSMIC-311[4] would help to reduce uncertainty around long-term OS estimates for cabozantinib and BSC. However, there are no further planned data-cuts of the trial beyond CCO2.

Table 6: Company's updated base case and EAG additional sensitivity analyses undertaken in 2L subgroup

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER	DM
Company's updated base case (including EAG's error correction†)								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£24,199**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA6: EAG preferred analysis (deterministic)								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£30,218**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
EA6: EAG preferred analysis (probabilistic)								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£31,015**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA1a: Exponential OS with treatment effect waning at 3 years								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£39,157**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA1b: Hybrid KM + exponential tail after 12 months, constant HR								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£31,084**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA1c: Hybrid KM + exponential tail after 12 months, BSC hazard rate in both groups								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£59,448**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA2a: COSMIC-311 utility value in progression-free state								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£33,840**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA2b: DECISION trial utility values								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£31,617**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-	-	
ASA3: AE QALY losses doubled								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£30,514**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-		
ASA4: ECG costs doubled								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£30,684**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-		
ASA5: CT scan costs for BSC removed								
Cabozantinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£30,203**	**1.2**
BSC	[redacted]	[redacted]	[redacted]	-	-	-		

EAG - External Assessment Group; 2L - second-line; LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; DM - decision modifier; OS - overall survival; KM - Kaplan-Meier; AE - adverse event; ECG - electrocardiogram; CT - computerised tomography; BSC - best supportive care

** Undiscounted*

† The company's TE model includes an error whereby the general population utility cap has been removed. This corrected analysis includes the general population utility cap.

4. References

1. Ipsen. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Company's technical engagement response form. Slough; 2023.
2. Ipsen. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Document B - company's evidence submission. Slough, UK; 2022.
3. Tappenden P, Carroll C, Ren K, Navega Biz A, Ren S, Clowes M. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Sheffield; 2023.
4. Brose MS, Robinson B, Sherman SI, Krajewska J, Lin CC, Vaisman F*, et al.* Cabozantinib for radioiodine-refractory differentiated thyroid cancer (COSMIC-311): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncology* 2021;22:1125-38.
5. Fordham BA, Kerr C, Freitas HM, Lloyd AJ, Johnston K, Pelletier CL. Health state utility valuation in radioactive iodine-refractory differentiated thyroid cancer. *Patient Preference and Adherence* 2015;9:1561-72.
6. European Medicines Agency. Summary of product characteristics - cabometyx. London, UK; 2022.
7. Ipsen. Cabozantinib for previously treated differentiated thyroid cancer unsuitable for or refractory to radioactive iodine [ID4046]. Company's response to clarification questions from the EAG. Slough, UK; 2022.
8. Che Z, Green N, Baio G. Blended survival curves: A new approach to extrapolation for time-to-event outcomes from clinical trials in health technology assessment. *Medical Decision Making [Epub ahead of print]* 2022.
9. Hernández-Alava M, Pudney S, Wailoo A. Estimating EQ-5D-3L by age and sex for the UK. Sheffield, UK; 2022.
10. National Institute for Health and Care Excellence. Technology Appraisal 516 - Cabozantinib for treating medullary thyroid cancer. London, UK; 2018.
11. National Institute for Health and Care Excellence. TA550 - Vandetanib for treating medullary thyroid cancer. London; 2018.
12. National Institute for Health and Care Excellence. TA742 - Selpercatinib for treating advanced thyroid cancer with RET alterations. London, UK; 2021.
13. Exelixis Inc. XL184-311 Clinical study report (COSMIC-311, CCO1). 2020.
14. Exelixis Inc. XL184-311 Clinical study report - Addendum 2 (COSMIC-311, CCO2); 2021.

Appendix 1: Cost-effectiveness results including QALY weighting

Table 7: Company's TE base case and EAG additional sensitivity analyses undertaken in 2L subgroup

Scenario	DM	ICER excluding QALY weighting	ICER including QALY weighting
Company's updated base case analysis, 2L subgroup (excluding EAG's error correction)	1.2	£23,050	£19,208
Company's updated base case (including EAG's error correction)	1.2	£24,199	£20,166
EA6: EAG preferred analysis (deterministic)	1.2	£30,218	£25,181
EA6: EAG preferred analysis (probabilistic)	1.2	£31,015	£25,878
ASA1a: Exponential OS with treatment effect waning at 3 years	1.2	£39,157	£32,630
ASA1b: Hybrid KM + exponential tail after 12 months, constant HR	1.2	£31,084	£25,904
ASA1c: Hybrid KM + exponential tail after 12 months, BSC hazard rate in both groups	1.2	£59,448	£49,540
ASA2a: COSMIC-311 utility value in progression-free state	1.2	£33,840	£28,200
ASA2b: DECISION trial utility values	1.2	£31,617	£26,348
ASA3: AE QALY losses doubled	1.2	£30,514	£25,429
ASA4: ECG costs doubled	1.2	£30,684	£25,570
ASA5: CT scan costs for BSC removed	1.2	£30,203	£25,169

EAG - External Assessment Group; 2L - second-line; DM - decision modifier; ICER - incremental cost-effectiveness ratio; QALY - quality-adjusted life year; EA - exploratory analysis; ASA - additional sensitivity analysis; KM - Kaplan-Meier; HR - hazard ratio; AE - adverse event; ECG - electrocardiogram; CT - computerised tomography; BSC - best supportive care