NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Committee Papers

ISBN: 978-1-4731-4841-3

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Cabozantinib for previously treated advanced hepatocellular carcinoma (review of TA582) [ID3917]

Contents:

The following documents are made available to consultees and commentators:

The **final scope** and **final stakeholder list** are available on the NICE website.

1. **Company submission** from Ipsen Limited UK
   a. Company FTA submission
   b. Company STA submission

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submissions** from:
   a. British Liver Trust
   b. British Association for the Study of the Liver (BASL)
   c. National Cancer Research Institute-Association of Cancer Physicians-Royal College of Physicians-Royal College of Radiologists

4. **Expert personal perspectives** from:
   a. Professor Tim Meyer, Consultant Medical Oncologist – clinical expert, nominated by Ipsen Limited UK
   b. Dr Richard Hubner, Consultant in Medical Oncology – clinical expert, nominated by the British Association for the Study of the Liver (BASL)

      Vanessa Hebditch – patient expert, nominated by the British Liver Trust (see item 3a.)

5. **Evidence Review Group report** prepared by the School of Health and Related Research (ScHARR)
   a. ERG report
   b. ERG addendum

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

Fast track appraisal: cost-comparison case

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Document B

Company evidence submission

February 2022

File name	Version	Contains confidential information	Date
ID3917_Cabozantinib Cost Comparison_DocB_AICCIC_Final.docx	**1.0**	**Yes**	**07/02/2022**

Contents

List of Tables

List of figures

Abbreviations

AEs	Adverse events
AFP	Alpha-fetoprotein
AIC	Akaike information criterion
ALT	Alanine aminotransferase
ANC	Absolute neutrophil count
AST	Aspartate aminotransferase
ATA	Adequate tumour assessments
BASL	British Association for the Study of the Liver
BCLC	Barcelona Clinic Liver Cancer
BIC	Bayesian information criterion
BOR	Best overall response
BNF	British National Formulary
BSC	Best supportive care
BSG	British Society of Gastroenterology
CI	Confidence interval
CR	Complete response
CRF	Case report form
CT	Computed tomography
CTCAE	Common terminology criteria for adverse events
EASL	European Association for the Study of Liver Diseases
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EHS	Extrahepatic spread
EMA	European Medicines Agency
EPAR	European public assessment report
EQ-5D	European Quality of Life Five Dimension
EQ5D-5L	Euroqol 5-dimension, 5-level health questionnaire
ESMO	European Society for Medical Oncology
FTA	Fast track appraisal
HbA1c	Haemoglobin A1c
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HCC	Hepatocellular carcinoma
HR	Hazard ratio
HRQoL	Health-related quality of life
IDMC	Independent data monitoring committee
ILCA	International Liver Cancer Association
IPD	Individual Patient Level Data
ITC	Indirect treatment comparisons
ITT	Intention to treat
KIT	Mast/stem cell growth factor
MAIC	Matching-adjusted indirect comparison
MET	Mesenchymal epithelial transition factor
MRI	Magnetic resonance imaging

MTC	Medullary thyroid cancer
MVI	Macroscopic vascular invasion
NASH	Non-alcoholic steatohepatitis
NCDFL	National Cancer Drug Fund List
NCRI	National Cancer Research Institute
NHS	National Health Service
NHSE	National Health Service England
NICE	The National Institute for Health and Care Excellence
NPACT	Non-protocol anticancer therapy
OS	Overall survival
OR	Odds ratio
ORR	Objective response rate
PD	Progressive/progressed disease
PD-1	Programmed cell death-1
PD-L1	Programmed cell death ligand 1
PFS	Progression-free survival
PH	Proportional Hazards
PLD	Patient level data
PR	Partial response
PS	Performance status
RCT	Randomised controlled trial
RECIST	Response evaluation criteria in solid tumours
RET	Rearranged during transfection
RTKs	Receptor tyrosine kinases
RWE	Real world evidence
SAEs	Serious adverse events
SD	Stable disease
SLR	Systematic literature review
SmPC	Summary of product characteristics
SoD	Sum of target lesion diameters
STA	Single technology appraisal
TA	Technology appraisal
TACE	Transarterial chemoembolisation
TIE	Angiopoietin
TKIs	Tyrosine kinase inhibitors
TTD	Time to treatment discontinuation
UK	United Kingdom
ULN	Upper limit of normal
UPCR	Urine protein/creatinine ratio
USA	United States of America
VEGFR	Vascular endothelial growth factor receptor
VEGFR2	Vascular endothelial growth factor receptor-2

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

B.1.1 Decision problem

Population

The marketing authorisation is: *"Cabozantinib is indicated as monotherapy for the treatment of hepatocellular carcinoma (HCC) in adults who have previously been treated with sorafenib."*

The population defined in the final scope is adults with advanced HCC who have had sorafenib. The submission covers the technology's full marketing authorisation for this indication.

Comparator

The manufacturer is proposing that the appraisal of cabozantinib be considered under the National Institute of Care and Excellence (NICE) Fast Track Appraisal (FTA) cost comparison process. The NICE guide to the technology appraisal (TA) process states that a cost comparison case can be made if a health technology is likely to provide similar or greater health benefits at similar or lower cost than technologies already recommended in published TA guidance for the same indication [1].

For cabozantinib the relevant comparator is regorafenib, as it is the only technology recommended in published NICE guidance for the same indication as cabozantinib. The wording of the regorafenib marketing authorisation is: *"Regorafenib is indicated as monotherapy for the treatment of adult patients with HCC who have been previously treated with sorafenib."*

Regorafenib is recommended by NICE (TA555) [2] as an option for treating advanced unresectable HCC in adults who have had sorafenib, only if:

- they have Child Pugh grade A liver impairment and an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1;
- the company provides it according to the commercial arrangement.

The NICE recommendation includes a restriction on the eligible patient population based on the degree of liver impairment and performance status. This is because the clinical trial evidence for regorafenib is based on advanced HCC patients that have been previously treated with sorafenib, and who have an ECOG performance status of 0 or 1 and Child-Pugh grade A liver impairment and not those who have more severe liver disease or a poorer performance status. Following the NICE approval of atezolizumab plus bevacizumab in first line (TA666) [3], regorafenib is now also being used in the third-line setting (further details regarding the treatment pathway are found in Section B.1.3). The positioning and use of regorafenib as the comparator for cabozantinib in clinical practice has been validated by clinical experts [4] treating eligible patients with drugs that are reimbursed according to the National Health Service England (NHSE) National Cancer Drugs Fund List (NCDFL) [5].

Ipsen wish to pursue the same positioning as the NICE recommendation for regorafenib as the clinical trial evidence is relatively limited for cabozantinib in people with advanced HCC with more severe liver disease or a poorer performance status.

It should be noted that best supportive care (BSC) is not a relevant comparator for a NICE FTA cost-comparison for cabozantinib, as the comparator can only be technologies already recommended in published technology appraisal guidance and/or treatment guidelines for the same indication.

Several analyses were conducted to provide evidence to support the comparative effectiveness of cabozantinib versus regorafenib, which consistently support similar or greater efficacy of cabozantinib versus regorafenib. A summary of the evidence includes the following:

- Indirect treatment comparisons (ITCs) using well-accepted and validated methodologies were conducted. The findings show no clear trend in ITC results in favour of cabozantinib versus regorafenib, but it can be concluded that cabozantinib is at least similar in clinical effectiveness to regorafenib and this conclusion is further supported from real world evidence (RWE) findings. The results of these analyses and RWE are described in greater detail in Section 3.10;

- Cabozantinib and regorafenib belong to same drug class of tyrosine kinase inhibitors (TKIs). They inhibit multiple receptor tyrosine kinases (RTKs) implicated in tumour growth, metastasis, and angiogenesis, including vascular endothelial growth factor receptor (VEGFR), endothelial-specific Angiopoietin receptor (TIE-2), mast/stem cell growth factor (KIT) and rearranged during transfection receptor (RET). The safety profile of cabozantinib is generally similar to that of other VEGFR-targeting TKIs. The results of adverse event comparisons are described in greater detail in Section 3.10;
- Clinical experts were consulted in an advisory board conducted by the manufacturer [4]. The clinical experts believe that the clinical effectiveness of cabozantinib and regorafenib are broadly equivalent. This is also reflected in the responses from the British Association for the Study of the Liver (BASL)/HCC-UK, British Society of Gastroenterology (BSG) and the National Cancer Research Institute (NCRI) Hepatobiliary Working Group, to the NICE scoping consultation for this topic; the NCRI Hepatobiliary Working Group also felt the FTA cost-comparison route was also appropriate [6].

To fulfil the criteria of "similar or lower costs", [redacted] As previously mentioned, regorafenib is the only approved NICE therapy in this indication, which is further supported by the findings of the clinical expert advisory board [4].
Since regorafenib fulfils all of the above criteria for a comparator in a FTA, a cost-comparison is considered an applicable method of economic analysis. The decision problem addressed by this submission is summarised in Table 1.

.

Table 1:The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with advanced hepatocellular carcinoma who have had sorafenib.	Adults with advanced hepatocellular carcinoma who have had sorafenib.	N/A
Comparator(s)	Regorafenib Best supportive care (BSC)	Regorafenib	BSC is not a relevant comparator in a cost-comparison case as the comparator can only be technologies already recommended in published technology appraisal guidance and/or treatment guidelines for the same indication.
Outcomes	• Overall survival • Progression-free survival • Response rates • Time to treatment discontinuation • Adverse effects of treatment • Health-related quality of life	• Overall survival • Progression-free survival • Response rates • Time to treatment discontinuation • Adverse effects of treatment • Health-related quality of life	The published literature for regorafenib does not present time to treatment discontinuation, limiting a comparison using this outcome
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. If the technology is likely to provide similar or greater health benefits at similar or lower cost than technologies recommended in published NICE technology appraisal guidance for the same indication, a cost-comparison may be carried out. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from a NHS and Personal Social Services perspective.	As per final scope.	N/A

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	The availability of any patient access schemes for the intervention, comparator or subsequent treatment technologies will be taken into account.		
Subgroups to be considered	None specified.	None specified.	N/A
Special considerations including issues related to equity or equality	N/A	No equity or equality issues for consideration	N/A

Abbreviations: NHS, National Health Service; NICE, The National Institute for Health and Care Excellence.

B.1.2 Description of the technology being appraised

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

Table 2: Technology being appraised

UK approved name and brand name	Cabozantinib (Cabometyx®)
Mechanism of action	Cabozantinib is an oral multi-targeted inhibitor of receptor tyrosine kinases (RTKs) that potently inhibits several RTKs known to influence tumour growth, metastasis and angiogenesis, including MET, VEGFR2 and AXL [7] . Treatment with cabozantinib results in anti-angiogenic effects in xenograft tumours, with disruption of the vasculature beginning within 24 hours after administration and is associated with pro-apoptotic effects leading to significant tumour growth inhibition or tumour regression in multiple tumour models including HCC, medullary thyroid cancer (MTC), breast cancer, lung carcinoma, glioblastoma and renal cell carcinoma [8-11]. The broad clinical activity of cabozantinib was demonstrated in a Phase I trial, in which tumour regression was observed in multiple tumour types [12] and these early findings were confirmed in a phase II randomised discontinuation trial (XL184-203 RDT) conducted in 9 tumour types, including HCC [13].
Marketing authorisation/CE mark status	An application for the marketing authorisation for cabozantinib in this indication was submitted to the European Medicines Agency (EMA) on 31 March 2018. The marketing authorisation process for the United Kingdom (UK) was centralised through the EMA at that time. The EMA granted marketing authorisation for cabozantinib, as monotherapy for the treatment of HCC in adults who have previously been treated with sorafenib, in November 2018.
Indications and any restriction(s) as described in the summary of product characteristics	Cabozantinib is indicated as monotherapy for the treatment of HCC in adults who have previously been treated with sorafenib. See Appendix C for the Summary of Product Characteristics and European public assessment report (EPAR).
Method of administration and dosage	Oral administration: One 60mg tablet to be taken once daily. Management of suspected adverse drug reactions may require temporary treatment interruption and/or dose 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. If a patient misses a dose, the missed dose should not be taken if it is less than 12 hours before the next dose.
Additional tests or investigations	A biopsy is required to establish histological or cytological diagnosis of HCC. Radiographic tumour assessment was performed every eight weeks using computed tomography or magnetic resonance imaging to assess disease progression in the pivotal trial. It is also recommended to monitor biochemical and metabolic parameters during treatment. This monitoring would likely be carried out as part of the routine management of advanced HCC.
List price and average cost of a course of treatment	£5,143.00 per 30 tablet pack. [14] Average cost per course of treatment equal to [redacted]

	[redacted].
Patient access scheme (if applicable)	A confidential simple patient access scheme is available. The pack price under this scheme is [redacted] (a [redacted]% discount to the list price). Under this scheme the average cost of a course of treatment (based on a [redacted]) is [redacted].

Abbreviations: EMA, European Medicines Agency; EPAR, European public assessment report; HCC, hepatocellular carcinoma; MET, mesenchymal epithelial transition factor; MTC, medullary thyroid cancer; RTKs, receptor tyrosine kinases, UK: United Kingdom; VEGFR, vascular endothelial growth factor receptor.

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

Disease overview

HCC is a primary hepatic cancer derived from well-differentiated hepatocytes [15]. It is the most common histologic subtype of liver cancer [16], accounting for approximately 80% of all liver cancers cases (estimates range from 70 to 90%) [17-21]. HCC occurs predominantly in patients with underlying chronic liver disease and cirrhosis; typically associated with viral hepatitis, excessive alcohol consumption, non-alcoholic steatohepatitis and haemochromatosis [17].

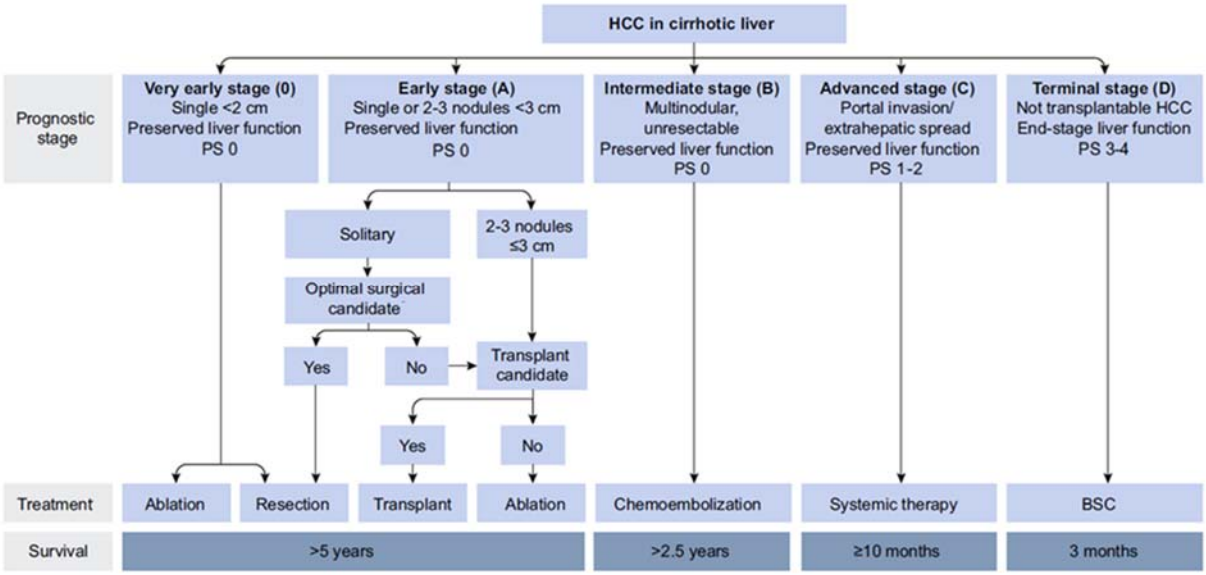
In the United Kingdom (UK), HCC is amongst cancers with the most rapid rate of growth both in incidence and mortality in last few decades [22]. There are 6,214 new cases of liver cancer each year in UK (2016-2018) with 66% cases in males [23]. The European Age-Standardised incidence rate in the UK (2016-2018) was 10 per 100,000 population; with significantly higher rates in men (14.5 per 100,000) compared to women (6.2 per 100,000) [23]. Over the last decade (between 2006-2008 and 2016-2018), the liver cancer age-standardised incidence rate increased by 45% in the UK [23]. It is projected to rise by 38% between 2014 and 2035, to 15 per 100,000 people by 2035 [23]. Approximately 5,600 deaths are caused by liver cancer in the UK every year, accounting for 3% of all cancer deaths in 2018 [24]. The liver cancer age-standardised mortality rates increased by 48% in the UK over the last decade [24], which is projected to rise by 58 between 2014 and 2035, to 16 deaths per 100,000 people by 2035 [24].

The overall prognosis for HCC depends on the severity of underlying liver dysfunction and the prognosis remains poor due to rapid disease progression and low survival

rates. The age-standardised net survival rate at 1 year is 38.1%, and at 5 years is 12.7% in UK [24].

The Barcelona Clinic Liver Cancer (BCLC) staging system is widely used in the UK and is endorsed by the European Association for the Study of Liver Diseases (EASL) [25], the European Society for Medical Oncology (ESMO) [26] and the American Association for the Study of Liver Diseases [27]. The BCLC classification divides HCC patients into five stages (0, A, B, C and D). The Child-Pugh status, measuring severity of cirrhosis involves five clinical measures and scoring them between 1 and 3. The sum of all five measures gives Child-Pugh score which leads to a classification of Child-Pugh A, B or C, with C being the most severe [28]. The classification of HCC is presented in Figure 1 and Table 3.

Figure 1: Barcelona Clinic Liver Cancer (BCLC) staging system and treatment strategy (EASL Guidelines)



Abbreviations: HCC, hepatocellular carcinoma; PS, performance status.
Source: EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma [29].

Table 3: Staging of HCC using BCLC classification

BCLC Staging	Tumour status	ECOG performance status	Liver function (Child-Pugh)
Stage 0 (Very early HCC)	Singe tumour <2cm in diameter without vascular invasion/satellites	0	Well preserved function Child-Pugh A
Stage A (Early HCC)	Single tumours >2cm or up to 3 nodules <3 cm in diameter	0	Child-Pugh A or B
Stage B (Intermediate HCC)	Multinodular asymptomatic tumours without an invasive pattern	0	Child-Pugh A or B

BCLC Staging	Tumour status	ECOG performance status	Liver function (Child-Pugh)
Stage C (Advanced HCC)	Symptomatic tumours; macrovascular invasion (either segmental or portal invasion) or extrahepatic spread (lymph node involvement or metastases)	0-2*	Child-Pugh A or B
Stage D (End stage HCC)		3-4	Child-Pugh C

*ESMO guidelines describe Stage C (advanced HCC) with ECOG performance status of 0-2 [26]
Abbreviations: BCLC staging, Barcelona clinic liver cancer staging; ECOG, Eastern Cooperative Oncology Group; HCC, Hepatocellular carcinoma.

Treatment for HCC depends on the location and stage of the cancer, and status of liver functioning. Approximately 30-40% of HCC patients worldwide, who are diagnosed with very early or early disease (BCLC stage 0/A), are eligible for curative procedures, including surgery (hepatic resection or liver transplantation) or percutaneous ablation [27, 29-31]. Around half of patients with HCC undergoing resection have a relapse in less than 3 years [31].

Patients with intermediate-stage HCC (BCLC B), in whom liver function is preserved, may be candidates for transarterial chemoembolisation (TACE) [32]. Most patients are diagnosed in the advanced stages of the disease (BCLC stage C), when cirrhosis is present and surgery is rarely an option, with the disease considered incurable [33]. Without treatment, the median survival for BCLC stage C patients ranges between 4 and 8 months [34].

Current treatment pathway for advanced HCC

For patients with advanced HCC, treatment options include interventional procedures such as TACE (using doxorubicin or cisplatin) or selective internal radiation therapy, and external beam radiotherapy. Patients unresponsive to these therapies, or with metastatic disease, are treated with systemic therapies.

UK and European clinical practice guideline recommendations

The BSG guidelines for HCC in UK clinical practice were published in 2003, prior to sorafenib and regorafenib becoming available [35]. As this existing guideline is outdated, the UK clinical practice largely aligns with the NICE treatment pathway and European guidelines published by ESMO and EASL [26, 29, 36]. The EASL guidelines, published in 2018 prior to EMA's approval for cabozantinib in HCC,

recommended sorafenib and lenvatinib in first-line and regorafenib in second-line [29]. The 2021 EASL position paper complements the 2018 guidance, recommending sorafenib, atezolizumab plus bevacizumab, and lenvatinib as first-line treatments [37]. In second-line, post atezolizumab plus bevacizumab, the 2021 EASL position paper recommends multi targeted tyrosine kinase inhibitors (TKIs) and vascular endothelial growth factor receptor-2 (VEGF-2) TKIs [37]. In second-line post sorafenib and lenvatinib, regorafenib, cabozantinib and ramucirumab are recommended treatments [37].

The ESMO clinical practice guidelines [36], updated in March 2021, recommends atezolizumab plus bevacizumab, sorafenib and lenvatinib as first-line treatments. After treatment with sorafenib, the guidelines recommend cabozantinib, regorafenib and ramucirumab as 'standard' second-line treatments. As there is no evidence for any drug in particular, ESMO guidelines recommend that all the currently approved first- and second-line agents could be considered as second-line therapy post atezolizumab plus bevacizumab i.e. sorafenib, lenvatinib, cabozantinib, regorafenib and ramucirumab [36].

The 2020 International Liver Cancer Association (ILCA) guidelines also recommend atezolizumab plus bevacizumab as standard of care, with exception in patients for whom atezolizumab or bevacizumab are contraindicated (sorafenib and lenvatinib are recommended as alternative option) [38]. Although, there is no data to support one TKI over another, the ILCA guidelines suggest sorafenib, lenvatinib and cabozantinib, after treatment with atezolizumab plus bevacizumab in first-line [38]. The ILCA guidelines also supported the use of regorafenib (in patients who tolerated sorafenib) and ramucirumab (in patients with alpha fetoprotein (AFP) ≥400 ng/mL) and cabozantinib, as second-line treatment after first-line sorafenib [38]. The ILCA guidelines also suggest sorafenib, cabozantinib, regorafenib, and ramucirumab as second-line options if lenvatinib is used first-line, although there are no data to support this [38].

Atezolizumab plus bevacizumab is the first treatment to demonstrate a significant OS benefit compared with sorafenib and consequently the treatment landscape has changed with atezolizumab plus bevacizumab becoming the standard of care in first-line systemic therapy for advanced HCC in the UK [4]. For patients having treatment

with atezolizumab plus bevacizumab, the median progression-free survival (PFS) is only 6.8 months, raising the need to define options for second-line therapy [36, 38]. Drugs in the second-line setting have so far only been tested after sorafenib failure/intolerance and there are currently no phase III trial data to inform the choice of second-line therapy in HCC patients that received alternative front-line therapies. There is, however, a clear rationale for offering a multikinase inhibitor given the existing evidence for efficacy in first and second-line.

The EASL, EMSO and ILCA guidelines are summarised in Table 4.

Table 4: EASL, ESMO and ILCA summary of guidelines

	ESMO 2021	EASL 2021	ILCA 2020
First-line	Standard: • Atezolizumab plus bevacizumab Option: • Sorafenib • Lenvatinib	Atezolizumab plus bevacizumab If contraindications to atezolizumab plus bevacizumab: • Sorafenib • Lenvatinib	First choice: Atezolizumab plus bevacizumab Alternative: • Sorafenib • Lenvatinib
Second-line	Option post atezolizumab plus bevacizumab: • Cabozantinib • Sorafenib • Lenvatinib • Regorafenib (only in patients previously exposed to TKIs) • Ramucirumab (only in patients with an AFP level ≥400 ng/mL) Standard post-sorafenib: • Cabozantinib • Regorafenib (only in patients previously exposed to TKIs) • Ramucirumab (only in patients with an AFP level ≥400 ng/mL)	Post atezolizumab plus bevacizumab: • Multi-TKI and VEGFR2 inhibitor as per off-label availability Post-sorafenib or lenvatinib: • Cabozantinib • Regorafenib (in sorafenib-tolerant patients) • Ramucirumab (in patients with serum alpha--fetoprotein above 400 ng/ml)	Post atezolizumab plus bevacizumab: • Cabozantinib • Sorafenib • Lenvatinib Post-sorafenib: • Cabozantinib • Regorafenib (in patients who tolerated sorafenib) • Ramucirumab (If AFP ≥400 ng/mL) Post-lenvatinib first line: • Sorafenib • Cabozantinib • Ramucirumab (if AFP ≥400 ng/mL)

	ESMO 2021	EASL 2021	ILCA 2020
			• Regorafenib (in patients who tolerated sorafenib)

Abbreviations: AFP, alpha fetoprotein; EASL, European Association for the Study of Liver Diseases; ESMO, European Society for Medical Oncology; ILCA, International Liver Cancer Association; TKI, tyrosine kinase inhibitor
Source: EASL 2021 [37], ESMO 2021 [36], ILCA 2020 [38]

NICE recommendations for first-line systemic treatment of advanced and unresectable HCC

NICE has recommended atezolizumab plus bevacizumab combination, sorafenib, and lenvatinib as first-line systemic therapies for adult patients with advanced and unresectable HCC (TA666 [3], TA474 [39], TA551 [40]). NICE guidance based on these technology appraisals are presented in Table 5 and Figure 2 [4].

NICE recommendations for second and later-line systemic treatment of advanced and unresectable HCC

For patients who have had sorafenib, only one treatment option, regorafenib, currently exists and is the standard of care in the UK practice following treatment with sorafenib. NICE has recommended regorafenib for patients who have had sorafenib, only if they had Child-Pugh grade A liver impairment and an ECOG performance status of 0 or 1 (TA555) [2]. The clinical evidence for regorafenib is based on the RESORCE trial, which studied regorafenib as a second-line treatment option for patients receiving and tolerating sorafenib in the first-line setting. NICE guidance based on this technology appraisal is presented in Table 5 and Figure 2 [4]. Ramucirumab is not currently approved by NICE, and currently, there is no ongoing NICE appraisal for ramucirumab. UK clinical experts have also confirmed to Ipsen that ramucirumab is not used in UK clinical practice [4].

Table 5: Summary of NICE technology appraisals related to HCC

NICE Technology Appraisals	Date
Atezolizumab with bevacizumab for advanced or unresectable hepatocellular carcinoma (TA666) – *Atezolizumab plus bevacizumab is recommended as an option for treating advanced or unresectable hepatocellular carcinoma (HCC) in adults who have not had previous systemic treatment, only if they have Child-Pugh grade A liver impairment and an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1, and the company provides it according to the commercial arrangement.*	December 2020 NICE TA666

NICE Technology Appraisals	Date
Regorafenib for previously treated unresectable hepatocellular carcinoma (TA555) – *Regorafenib is recommended as an option for treating advanced unresectable hepatocellular carcinoma in adults who have had sorafenib, only if they have Child-Pugh grade A liver impairment and an ECOG performance status of 0 or 1, and the company provides it according to the commercial arrangement.*	January 2019 NICE TA555 (replaces TA514)
Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma (TA551) – *Lenvatinib is recommended as an option for untreated, advanced, unresectable hepatocellular carcinoma in adults, only if patients have Child-Pugh grade A liver impairment and an ECOG performance status of 0 or 1, and the company provides lenvatinib within the agreed commercial arrangement.*	December 2018 NICE TA551
Sorafenib for the treatment of advanced hepatocellular carcinoma (TA474) – *Sorafenib is recommended as an option for treating advanced hepatocellular carcinoma only for people with Child-Pugh grade A liver impairment, only if the company provides sorafenib within the agreed commercial access arrangement.*	September 2017 NICE TA474 (replaces TA189)

Abbreviations: ECOG, Eastern Cooperative Oncology Group; HCC, hepatocellular carcinoma; NICE, The National Institute for Health and Care Excellence; TA, technology appraisal.

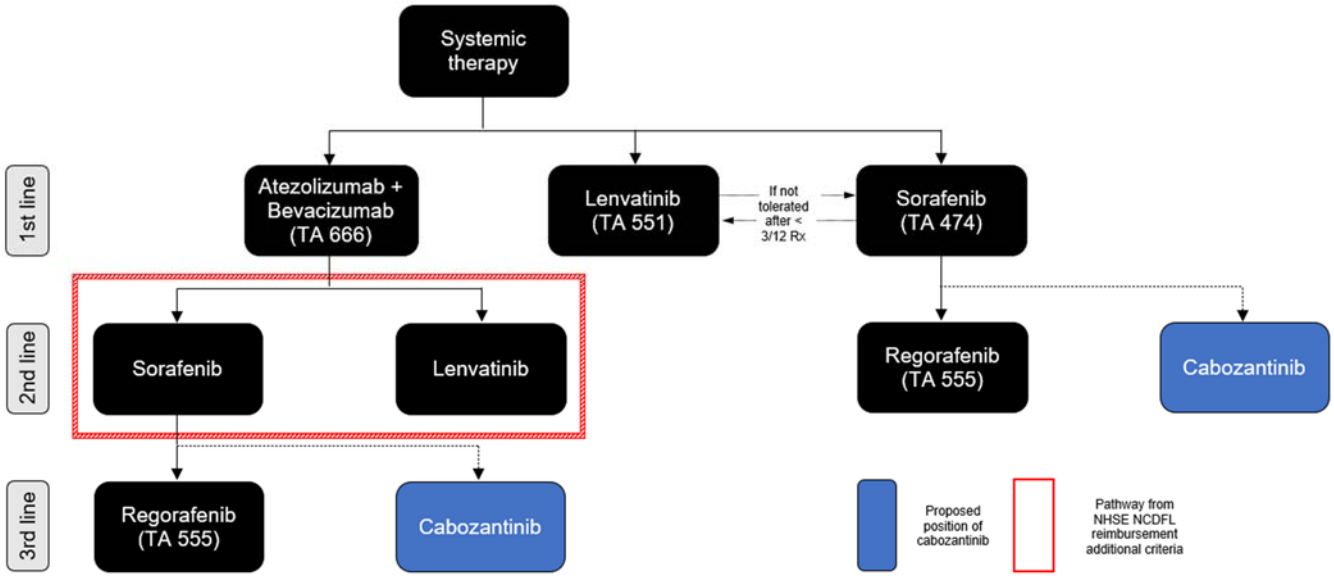
NHS England National Cancer Drugs Fund List (NHSE NCDFL)

The actual position of NICE recommended medicines that are reimbursed by NHSE NCDFL [5] is slightly different to the wording of the NICE recommendations. For lenvatinib and sorafenib use in second-line (despite all the evidence only being for their use in a first-line setting), there are additional specific criteria applied to account for the changing landscape, as atezolizumab plus bevacizumab becomes the standard of care in first-line systemic therapy in advanced HCC. These additional criteria, that are outside of NICE recommendations include:

- Ability to receive lenvatinib or sorafenib if the patient has received atezolizumab plus bevacizumab as first-line treatment;
- Ability to switch from lenvatinib to sorafenib (and vice versa) in the first-line setting if patient has had to discontinue treatment within 3 months of starting the drug and solely because of toxicity.

Thus regorafenib is currently prescribed and reimbursed post sorafenib by NHSE in either a second-line setting where sorafenib has been prescribed first-line or in a third-line setting (despite the lack of evidence demonstrating its efficacy beyond the second-line treatment setting) for patients previously treated with atezolizumab plus bevacizumab followed by sorafenib. There is no NICE approved recommendation for second-line treatment following first-line treatment with lenvatinib, although clinical experts would welcome a treatment option in this setting [4]. The positioning of these treatments has been confirmed by UK clinical experts [4] and is summarised in Figure 2 below.

Figure 2: Current systemic therapy treatment pathway in UK clinical practice as per NICE and NHSE NCDFL recommendations



Abbreviations: NCDFL, National Cancer Drug Fund List; NHSE, National Health Service England; NICE, National Institute for Health and Care Excellence; Rx, prescription.
Source: Clinical experts' opinion [4], NHSE NCDFL [5].

Positioning of cabozantinib

It is proposed that cabozantinib is positioned where regorafenib is currently used in practice as shown in Figure 2.

It can be argued that the evidence base and generalisability of cabozantinib for the UK advanced HCC population is greater than that of regorafenib in its current position for the following reasons:

- The pivotal clinical trial (CELESTIAL) [7] for cabozantinib had broader inclusion criteria than that of the regorafenib pivotal trial (RESORCE) [41] as it included:
    - Both second and third-line patients (28% of trial patients were receiving third-line therapy) whilst the RESORCE trial only included patients who had received sorafenib first-line only i.e. the regorafenib population were pure second-line;
    - Patients intolerant to sorafenib. The CELESTIAL trial included patients who had disease progression on sorafenib irrespective of whether they had tolerated sorafenib or not, unlike the RESORCE trial where patients who had disease progression on sorafenib had to have tolerated sorafenib (≥400 mg daily for at least 20 of the 28 days before discontinuation);
    - Additionally, compared to the RESORCE trial patients in the CELESTIAL trial were more likely to be white (56% versus 36%), and less likely to be in the Asia geographical region (25% versus 38%).

This makes cabozantinib a more relevant treatment option than regorafenib in practice when taking into account the current NICE and NHSE NCDFL recommendations.

Due to differences in the CELESTIAL and RESORCE trials designs, no superiority claim is made for cabozantinib in this submission. However, cabozantinib is currently the only therapy developed for HCC that inhibits the MET and AXL receptors (in addition to VEGFR 1, 2 and 3), and thereby provides additional inhibitory effects beyond that of currently approved TKIs [9]. Due to this unique molecular pathway, cabozantinib may be able to break TKI resistance established in the first-line of treatment [42-44]. Therefore, cabozantinib has a biologically plausible rationale to treat patients who are resistant to sorafenib [7].

This submission aims to demonstrate that cabozantinib does fulfil the FTA cost-comparison criteria by being a health technology that is likely to provide similar or greater health benefits at similar or lower cost than technologies already recommended (i.e., regorafenib) in published technology appraisal guidance for the same indication. In addition, it aims to demonstrate that the evidence base for cabozantinib is more generalisable to UK practice and thus offers an additional treatment option for UK patients with advanced HCC, where systemic treatment

options are limited and the prognosis remains poor as they continue to progress rapidly and have a short overall survival (OS) of 8 to 11 months [25, 41].

B.1.4 Equality considerations

No equality issues related to the use of cabozantinib have been identified.

B.2 Key drivers of the cost-effectiveness of the comparator(s)

B.2.1 Clinical outcomes and measures

Only NICE TA555 is relevant to this submission as it addresses the same decision problem concerning the treatment of patients with advanced HCC who have previously been treated with sorafenib [2].

Clinical efficacy

Clinical trials investigating the use of traditional cytotoxic agents considered tumour response as the necessary primary endpoint. Molecular targeted therapies have shown improved survival with no measurable change in tumour size. As a result, time-to-event endpoints are preferred as indicators of treatment efficacy for molecular targeted therapies, rather than decreases in tumour size [25, 45].

OS is the primary endpoint showing least investigator bias [45]. OS captures the time from randomisation until death due to any cause. In TA555 [2] the median follow-up was 7.0 months with 40% of the regorafenib arm alive at the end of the follow-up period.

PFS, providing evidence of radiological progression, is recommended as a secondary endpoint in pivotal phase III HCC trials [45]. It is defined as the time from randomisation to the occurrence of disease progression or death from any cause, whichever occurred first. However, it is sometimes considered unreliable in HCC as death resulting from the natural history of cirrhosis might confound the detection of potential clinical benefit [45]. Another secondary efficacy endpoint used in HCC trials was analysis of the investigator-determined objective response rate (ORR).

Measuring health-related quality of life (HRQoL) for patients with HCC is a challenging outcome as impaired HRQoL may be a consequence of the natural history of underlying liver disease and not of tumour progression [45].

The relevance of all these endpoints is discussed for cabozantinib in Section B.3.

Other key clinical outcomes: Adverse events and discontinuation rates

In addition to clinical response, the incidence of adverse events (AEs) and treatment discontinuation, as well as their impact for decision-making, have been frequently discussed during NICE committee meetings.

The relatively high frequency of AEs, even in the placebo group, reflects the high burden of advanced HCC and underlying liver disease in this patient population. The ability to maintain quality of life on treatment is an important driver for continuing treatment. Discontinuation of the treatment due to AEs is reported for the trial. Inclusion of AEs has limited impact on the cost-effectiveness analysis of the treatment despite the poor prognosis of the patient group.

Table 6: Clinical outcomes and measures appraised in published NICE STA guidance for the comparator(s)

	Outcome	Used in cost-effectiveness modelling	Committee's preferred assumptions	Uncertainties (if applicable)
NICE TA555 (Regorafenib)	Overall survival	Standard parametric models fitted to patient level data from the RESORCE trial. Dependent log normal curves were used in the manufacturer's original base case	Committee preferred independent Weibull curves for extrapolating overall survival	• Committee recognised that Weibull curves were associated with significant uncertainty due to the immaturity of the data
	Progression-free survival	Kaplan-Meier data from RESORCE used directly for progression-free survival curve	Committee agreed that the data from RESORCE represented the full pattern of progression	
	HRQoL	EQ-5D data, use of EQ-5D questionnaire	• High utility values used in the model did not seem clinically plausible in patients with progressed HCC and was likely to have resulted in an underestimate of the ICERs • An EQ-5D questionnaire was completed on the first day of each treatment cycle, when a patient had not had treatment for a week, therefore, adverse events have not been fully captured	• Committee was concerned about face validity of the utility values collected using EQ-5D data because the utility decrement for progression appeared low for an advanced hepatocellular population with progressed disease

	Outcome	Used in cost-effectiveness modelling	Committee's preferred assumptions	Uncertainties (if applicable)
	Discontinuation rate	Discontinuation rates were applied to each health state using time on treatment data from RESORCE	• The committee concluded rate of treatment discontinuation in RESORCE is unlikely to represent NHS clinical practice • Adjusting for cost alone for 20% of people having treatment post progression was unreasonable • Fully extrapolating time to treatment discontinuation from RESORCE using standard parametric models fitted to individual patient level data.	• Number of people continuing treatment despite disease progression and the efficacy of treatment in these patients was uncertain • People would have less treatment in practice than in RESORCE, as they discontinue if disease progresses

Abbreviations: EQ-5D, European Quality of Life Five Dimension; HCC, hepatocellular carcinoma; HRQoL, Health-Related Quality of Life; NHS, National Health Service; TA, Technology Appraisal.
Source: TA555 [2].

B.2.2 Resource use assumptions

Resource use considered in the relevant NICE technology appraisal (TA555) is listed below:

- Drug acquisition:
    - This was informed by discontinuation rates mentioned in Table 6
- Disease health state-specific cost components:
    - Monitoring by medical staff including oncologists, hepatologists, gastroenterologists, specialised nurses. This also includes both routine laboratory monitoring tests and radiological tests;
    - Systemic medications, inpatient admissions, and outpatient care;
    - Hospitalisation due to advanced stage of HCC and prior treatments;
- Hospitalisation costs due to AEs and for terminal care of patients.

The clinical expert in TA555 explained that 80% of patients would stop treatment on progression and since there was a high number of people continuing treatment in the RESORCE trial, this would not be representative of clinical practice. Alternative scenarios explored different costs of post progression but ultimately the committee concluded that it was inappropriate to adjust only for cost and not health benefit. For regorafenib, the company used clinician surveys to estimate resource use associated with sorafenib and BSC. It was assumed that the sorafenib results would also apply to regorafenib. The committee was not convinced of the robustness of the surveys and noted the small number of clinicians involved and the variability in the clinicians' responses. Estimates from the 2007 and 2015 surveys were therefore pooled for health state resource use costs. Considering all committee assumptions, costs were calculated, using revised rates of hospitalisation and assuming wastage of medicine [2].

B.3 Clinical effectiveness

Cabozantinib significantly improved OS, PFS and the ORR compared with placebo, with a manageable safety profile.

Clinical efficacy and Safety

- Cabozantinib significantly extended OS in advanced HCC patients versus placebo: median OS 10.2 months (95% CI: 9.1, 12.0) for cabozantinib versus 8.0 months (95% CI: 6.8, 9.4) for placebo, with a hazard ratio (HR) for death: 0.76 (95% CI: 0.63, 0.92; P = 0.005) [7]
- In the subgroup analysis of patients previously treated with sorafenib only, cabozantinib provided an additional 4.1 months of median OS versus placebo (11.3 months for cabozantinib and 7.2 months for placebo). Risk of death was reduced by 30% in this population (stratified HR for death: 0.70; 95% CI: 0.55, 0.88) [7]
- Cabozantinib significantly improved PFS in advanced HCC patients: median PFS 5.2 months (95% CI: 4.0, 5.5) versus 1.9 months (95% CI: 1.9, 1.9) for placebo, with a HR for disease progression or death: 0.44; (95% CI: 0.36, 0.52; P<0.001) [7]
- In the subgroup analysis of patients previously treated with sorafenib only, cabozantinib provided an additional 3.6 months of median PFS (5.5 months for cabozantinib and 1.9 months for placebo; HR for disease progression or death: 0.40; 95% CI: 0.32, 0.50) [7]
- AEs were consistent with the known safety profile of cabozantinib [7]

B.3.1 Identification and selection of relevant studies

See Appendix D, Section 1.1 for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised.

B.3.2 List of relevant clinical effectiveness evidence

Clinical evidence to support the use of cabozantinib for the treatment of advanced HCC comprises a single randomised controlled trial (RCT) – the CELESTIAL trial (XL184-309; NCT01908426). A brief overview of this trial is provided in Table 7.

A systematic review of the literature did not identify any additional studies relevant to cabozantinib in advanced HCC.

Table 7: Clinical effectiveness evidence

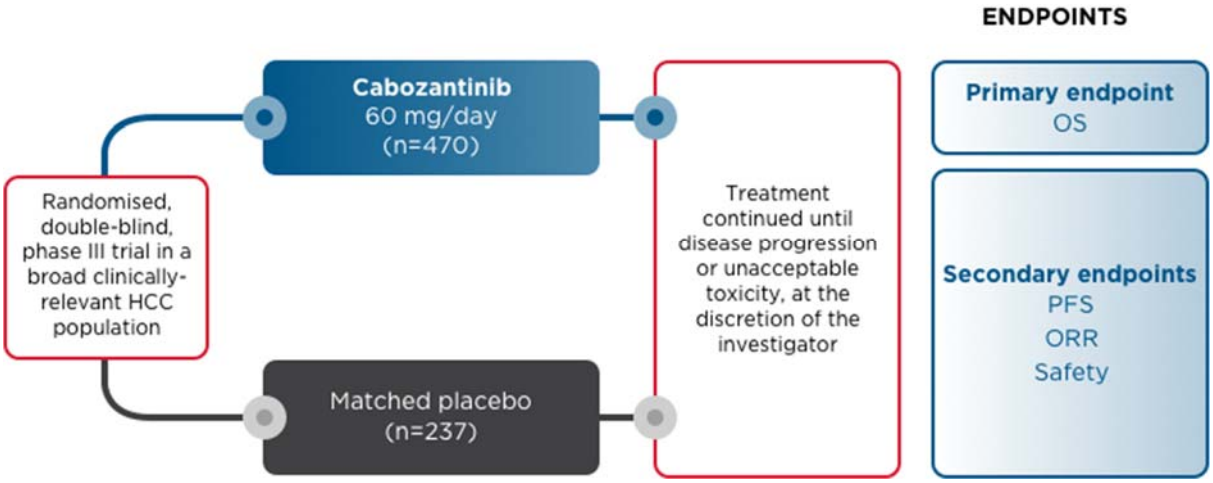
Study	CELESTIAL
Study Design	Randomised, double-blind, placebo-controlled, phase III
Population	Patients with previously treated advanced HCC
Intervention(s)	Oral cabozantinib 60 mg once daily plus best supportive care (BSC)
Comparators	Oral matched placebo once daily plus BSC
Does trial support application for marketing authorisation	Yes
If trial 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 (EQ5D-5L)
All other reported outcomes	• Pharmacokinetics

Abbreviations: AE, adverse events; BSC, best supportive care; EQ5D-5L, Health-related quality of life; HCC, hepatocellular carcinoma; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; TTD, time to treatment discontinuation.

B.3.3 Summary of methodology of the relevant clinical effectiveness evidence

CELESTIAL Trial: The CELESTIAL global phase III clinical trial tested the effects of cabozantinib compared with placebo in patients with advanced HCC who had already received treatment with sorafenib. The schematic design of the trials is depicted in Figure 3.

Figure 3: CELESTIAL trial design



Abbreviations: HCC, hepatocellular carcinoma; ORR, objective response rate; OS, overall survival; PFS, progression-free survival.
Source: Abou-Alfa et al., 2018 [7].

Outcome measures used in the economic model or specified in the scope

The relevant endpoints in the CELESTIAL trial along with details of when and how they were measured during the trial are summarised in Table 8 [46]. All endpoints and outcomes described were pre-specified, unless otherwise stated.

Table 8. Relevant endpoints and measures in the CELESTIAL trial

Endpoint	Definition	Timing and nature of assessment
Primary endpoint		
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
Secondary endpoints		
PFS	The date of randomisation to radiographical progression or death, whichever occurred first	Radiographical tumour assessment by the investigator (or radiologist) was based on Response Evaluation Criteria in Solid Tumours (RECIST 1.1) Computed tomography (CT)/magnetic resonance imaging (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)
ORR	The proportion of patients with a best overall response (BOR) of complete response (CR) or partial response (PR). CR or PR must be confirmed on a subsequent visit ≥28 days after the response was first observed	

Endpoint	Definition	Timing and nature of assessment
		Bone scans were performed at screening, 8 and 16 weeks after randomisation, and every 16 weeks in patients with documented bone lesions at screening or suspicion of bone metastasis during the trial 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
Exploratory endpoints		
HRQoL	Health status was measured using EQ-5D-5L 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	
Safety and tolerability	Safety assessments included the evaluation of AEs, serious AEs (SAEs), deaths, clinical laboratory tests (haematology, serum chemistry and urinalysis), physical examination, vital signs, ECOG PS, 12-lead electrocardiogram (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.	

Abbreviations: AE, adverse event; ECOG, Eastern Cooperative Oncology Group; HRQoL, health-related quality of life; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PS, performance status; TTD, time to treatment discontinuation.

A summary of the methodology of the Phase III CELESTIAL trial is presented in Table 9.

Table 9: Summary of trial methodology

Study	CELESTIAL TRIAL
Location	104 sites across 19 countries (Australia, Belgium, Canada, France, Germany, Hong Kong, Ireland, Italy, Republic of Korea, Netherlands, New Zealand, Poland, Romania, Singapore, Spain, Taiwan, Turkey, UK and USA)
Trial Design	Phase III, randomised, double-blind, controlled study of Cabozantinib versus Placebo in patients with HCC who have received prior Sorafenib
Eligibility criteria for participants	Inclusion criteria included: • Age ≥18 years of age on the day of consent • Histological or cytological diagnosis of HCC (previous biopsy results accepted) • Disease not amenable to curative treatment (e.g., transplant, surgery, radiofrequency ablation) • Received prior sorafenib • Progression following at least one prior systemic treatment for HCC • Recovery from toxicities related to any prior treatment to ≤Grade 1, unless the AEs were clinically non-significant and/or stable with supportive therapy • Eastern Cooperative Oncology Group (ECOG) performance status (PS): 0 or 1 at screening • Adequate haematological function, i.e., meeting the following laboratory criteria ≤7 days prior to randomisation: • Absolute neutrophil count (ANC): ≥1200/mm3 (≥1.2×109/L) • Platelets: ≥60,000/mm3 (≥60×109/L) • Haemoglobin: ≥8 g/dL (≥80 g/L) • Adequate renal function, i.e., meeting the following laboratory criteria ≤7 days prior to randomisation: • Serum creatinine ≤1.5×upper limit of normal (ULN) or calculated creatinine clearance ≥40 mL/min using the Cockcroft-Gault equation • Urine protein/creatinine ratio (UPCR) ≤1 mg/mg (≤113.1 mg/mmol) or 24-hour urine protein <1 g • Child-Pugh status: A • Total bilirubin ≤2 mg/dL (≤34.2 µmol/L) ≤7 days prior to randomisation • Serum albumin ≥2.8 g/dL (≥28 g/L) ≤7 days prior to randomisation • Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) <5.0×ULN ≤7 days prior to randomisation • Haemoglobin A1c (HbA1c) ≤8% within 28 days prior to randomisation (if HbA1c results were unavailable: fasting serum glucose ≤160 mg/dL) • If have active HBV infection, receiving antiviral therapy according to the local standard of care • Be capable of understanding and complying with the protocol requirements and providing written consent • Sexually active fertile subjects and their partners must have agreed to use medically accepted barrier methods of contraception during the trial and for 4 months after the last dose of study drug

Study	CELESTIAL TRIAL
	• Women of childbearing potential (premenopausal women capable of becoming pregnant and women who were amenorrheic for ≥12 months possibly due to prior chemotherapy, anti-oestrogens, ovarian suppression, low body weight or other reasons) must not have been pregnant at screening
Settings and location where the data were collected	The CELESTIAL trial was conducted in the secondary care setting in 19 countries: • Europe: Belgium, France, Germany, Ireland, Italy, Poland, Romania, Spain, The Netherlands, Turkey and United Kingdom • North America (United States of America [USA] and Canada) • Australia and New Zealand • Asia: Hong Kong, Republic of Korea, Singapore and Taiwan
Trial drugs	• Experimental Arm: Cabozantinib 60 mg tablet once daily • Comparator Arm: Matched placebo In addition, best supportive care was provided, based on the following general guidelines: • Analgesia and the management of AEs due to analgesia • Treatment of liver decompensation in patients with non-neoplastic liver disease • Antibiotics to treat infection, such as peritonitis and pneumonia • Provision of nutritional support and psychological support, including the management of depression and anxiety with medication and/or counselling • Transfusions to maintain haemoglobin levels, as clinically indicated (but not the use of erythroid growth factors).
Permitted and disallowed concomitant medication	The use of any of the following medications was permitted if required, during the trial: • Antiemetics and anti-diarrhoeal medications • Granulocyte colony-stimulating factors (except for prophylactic use before initial treatment with study drug) • Hormone replacement and short-term systemic steroid treatment • Low-doses of aspirin for cardio protection (per local guidelines), of warfarin (≤1 mg/day) and of low molecular-weight heparin • Antiviral therapy for active HBV infection. The use of any the following was not permitted in patients receiving study drug: • Any investigational agent or medical device • Any drug or herbal product specifically for the treatment of HCC • Therapeutic doses of oral anticoagulants (e.g., Warfarin [>1 mg/day] or warfarin-related agents, thrombin or factor Xa inhibitors) or antiplatelet agents (e.g., Clopidogrel); interferon • Liver-directed local anticancer therapy or systemic anti-tumour therapies • Erythropoietic-stimulating agents (e.g., Epoetin alfa and darbepoetin alfa)

Study	CELESTIAL TRIAL
	Palliative external radiation to bone metastasis or skin/subcutaneous metastasis was permitted during the trial but was discouraged unless medically unavoidable.
Primary outcome	• Overall Survival (OS) [Time Frame: Up to 45 months]
Secondary outcomes	• Progression-Free Survival (PFS) [Time Frame: Up to 45 months] • Objective Response Rate (ORR) [Time Frame: ORR is measured by radiologic assessment every 8 weeks after randomisation until disease progression or discontinuation of study treatment (up to 45 months)]
Exploratory endpoints	• HRQoL using EQ-5DL questionnaire • Safety and tolerability: evaluation of AEs, serious AEs (SAEs), deaths, clinical laboratory tests (haematology, serum chemistry and urinalysis), physical examination, vital signs, ECOG PS, 12-lead electrocardiogram (ECG) and the TTD in months (date of decision to discontinue study drug – date of first dose +1)/30.4375

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; ECOG, Eastern Cooperative Oncology Group; HbA1c, haemoglobin A1c; HBV, hepatitis B infection; HCC, hepatocellular carcinoma; ORR, objective Response Rate; OS, overall Survival; PFS, progression-free survival; PS, performance status; ULN, upper limit of normal; UPCR, urine protein/creatinine ratio; USA, United States of America.
Source: Exelixis, 2018 [46].

B.3.3.1 Patient baseline demographics and disease characteristics

The intention to treat (ITT) population included all patients randomised to receive study drug prior to the cut-off date for the second interim analysis, i.e., 1 June 2017, regardless of whether they received any/the correct study drug [7]. The ITT population comprised 470 patients in the cabozantinib group and 237 patients in the placebo group. Demographic and baseline characteristics in the ITT population were well balanced between the treatment groups. Overall, almost half of the study population were ≥65 years of age (49%) and 82% were male. Most patients were White (56%) or Asian (34%). ECOG performance status (PS) was 0 in 53% of patients and 1 in 47% of patients; a single patient in the cabozantinib group had an ECOG PS of 1 at screening and 2 at baseline [46].

Stratification factors in the ITT population were also balanced between the treatment groups (Table 10). The stratification factors consisted of the following:

- Etiology of disease (hepatitis B virus [HBV] (HBV [with or without hepatitis C virus (HCV)], HCV [without HBV], or Other)
- Geographic region (Asia, Other Regions)
- Presence of extrahepatic spread of disease and/or macrovascular invasion (Yes, No)

The majority of patients were enrolled in Europe or North America (72%), 25% were enrolled in Asia and 4% in Australia/New Zealand. HBV [with or without HCV] was present in 38% of patients, 21% had HCV (without HBV) and 40% had HCC of another aetiology. Most patients (78%) had extrahepatic disease spread and/or macrovascular invasion [46].

Table 10: Baseline characteristics of patients in the CELESTIAL trial

Study	CELESTIAL Trial	
Baseline patient and disease characteristics	**Cabozantinib (n=470)**	**Placebo (n=237)**
Age, years, Median (range)	64 (22, 86)	64 (24, 86)
Sex, n (%) **Male**	379 (81)	202 (85)

Study	CELESTIAL Trial	
Baseline patient and disease characteristics	**Cabozantinib (n=470)**	**Placebo (n=237)**
Race, n (%)		
White	264 (56)	130 (55)
Asian	159 (34)	82 (35)
Black or African American	8 (2)	11 (5)
Other	8 (2)	2 (1)
Not reported	31 (7)	12 (5)
Geographic region, n (%)		
Europe	231 (49)	108 (46)
Asia	116 (25)	59 (25)
North America (USA/Canada), n (%)	108 (23)	59 (25)
Australia/New Zealand	15 (3)	11 (5)
ECOG PS, n (%)		
0 (normal activity, asymptomatic)	245 (52)	131 (55)
1 (fully ambulatory, symptomatic)	224 (48)	106 (45)
2 (in bed <50% of time, ambulatory and capable of self-care but not work activities)	1 (<1)	0
Aetiology at baseline, according to the CRF, n (%)		
Dual HBV and HCV	8 (2)	4 (2)
HBV	178 (38)	89 (38)
HCV	113 (24)	55 (23)
Alcohol related	112 (24)	39 (16)
NASH	43 (9)	23 (10)
Other/unknown	99 (21)	63 (27)
Child-Pugh A status, according to the CRF, n (%)		
A (score 5–6)	462 (98)	235 (99)
B (score 7–9)	7 (1)	2 (0.8)
Missing	1 (0.2)	0
Baseline disease, according to the CRF, n (%)		
Extrahepatic spread	369 (79)	182 (77)
Macrovascular invasion	129 (27)	81 (34)
AFP ≥400 ng/mL, n (%)	192 (41)	101 (43)
Prior systemic non-radiation anticancer regimens for advanced HCC, n (%)		
0	3 (0.6)	0
1	335 (71)	174 (73)
2	130 (28)	62 (26)
≥3	2 (0.4)	1 (0.4)
Median (range)	1 (0, 3)	1 (1, 3)
Duration of prior sorafenib for HCC, months, median (range)	5.32 (0.3, 70.0)	4.80 (0.2, 76.8)
<1month, n (%)	11 (2)	8 (3)
≥1 to <3 months, n (%)	117 (25)	54 (23)
≥3 to <6 months, n (%)	130 (28)	67 (28)
≥6 months, n (%)	211 (45)	108 (46)
Time from progression on sorafenib as most recent prior systemic agent, months, median (range)	n=322 1.61 (0.1, 28.3)	n=166 1.66 (0.2, 69.4)
Prior local liver-directed therapy (including transarterial chemoembolisation [TACE]), for HCC, n (%)	209 (44)	113 (48)
Prior TACE, for HCC, n (%)	203 (43)	111 (47)

Baseline was considered the last observation prior to randomisation; multiple aetiologies could be reported for each patient.

Abbreviations: AFP, alpha-fetoprotein; CRF, case report form; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; ITT, intention to treat; NASH, non-alcoholic steatohepatitis; PD-1, programmed cell death-1; PD-L1, programmed cell death ligand 1.

Source: Abou-Alfa et al., 2018 [7], Exelixis, 2018 [46].

CELESTIAL trial population compared with a typical UK population with HCC

The overall study population in the CELESTIAL trial were largely similar to a typical population of patients with advanced HCC in the UK, based on a retrospective national audit including data from 448 patients from 15 hospitals who received first-line systemic therapy with sorafenib for HCC (Table 11) [47].

Due to the inclusion criteria in the CELESTIAL trial, a higher proportion of patients participating in this study had an ECOG PS of 0 and more patients had Child-Pugh status A compared with a typical population of patients with HCC (Table 11). A higher proportion of patients in the CELESTIAL trial had extensive metastatic disease at baseline, with almost double the proportion of patients with extrahepatic spread. In addition, a higher proportion of patients participating in the CELESTIAL trial had HBV and/or HCV (Table 11).

Table 11. Baseline and disease characteristics of a typical population of patients with advanced HCC in the UK (based on observational data) and participants in the CELESTIAL trial

	Observational data (N=448)	CELESTIAL trial (overall)* (N=707)
Age, years **Median (range)**	68 (17.0, 89.0)	64.0 (22, 86)
Sex, n (%)		
Male	325 (72.5)	581 (82.2)
Missing	57 (12.7)	0
ECOG PS, n (%) **0** **1** **2** **3** **Missing**	 117 (26.1) 218 (48.7) 94 (21.0) 6 (1.3) 13 (2.9)	 376 (53.2)) 330 (46.7) 1 (0.1) † 0 0
Disease characteristics		
Child-Pugh status, n (%) **A** **B** **C** **Missing**	 343 (76.6) 72 (16.1) 2 (0.4) 31 (6.9)	 697 (98.6) 9 (1.3) 0 1 (0.1)
Presence of extrahepatic spread, n (%) **Yes** **Missing**	 172 (38.4) 7 (1.6)	 551 (77.9) –
Presence of micro/vascular invasion, n (%) **Yes** **Missing**	 91 (20.3) 196 (43.8)	 210 (29.7) –
AFP ≥400 ng/mL, n (%)** **Missing**	141 (31.5) 80 (17.9)	293 (41.4) 0
Aetiology of disease, n (%) **HBV** **HCV** **Alcohol related**	 55 (12.3) 70 (15.6) 110 (24.6)	 267 (37.8) 168 (23.8) 151 (21.4)
Previous local therapy, n (%)	190 (42.4)	324 (45.8)

*Intention to treat population, according to the case report form (CRF) in the CELESTIAL trial
†A patient in the cabozantinib group had an ECOG PS of 1 at screening and 2 at baseline
** AFP ≥400 ng/mL defines a poorer prognositc group
Abbreviations: AFP, alpha-fetoprotein; ECOG, Eastern Cooperative Oncology Group; HBV, hepatitis B virus; HCV, hepatitis C virus; N/A, not available; PS, performance status.
Source: Abou-Alfa et al., 2018 [7], Exelixis, 2018 [46], King et al., 2017 [47],

B.3.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

B.3.4.1 Analysis populations

All efficacy analyses were conducted using data from the ITT population [46]. The results from the second planned interim analysis are presented in this document. For the second interim analysis, the ITT population comprised all patients randomised to receive study drug as of the cut-off date for the second interim analysis, i.e., 1 June 2017, regardless of whether they received any/the correct study drug. (Table 12)

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 12. Analysis sets in the CELESTIAL trial

Analysis sets	Number of patients		
	Cabozantinib	Placebo	Total
ITT			
Overall population	**470**	**237**	**707**
Safety			
Overall population	**467***	**237**	**704**

*Three patients **d**id not receive study drug
Abbreviations: ITT, intention to treat.
Source: Abou-Alfa et al., 2018 [7].

As of 18 September 2017, 773 patients had been enrolled in the trial (target sample size 760) and enrolment was closed [7].

B.3.4.2 Statistical analysis

An overview of the primary statistical analyses in the CELESTIAL trial is provided in Table 13 [46].

Sensitivity analyses

In addition to the primary analysis of PFS (PFS1), sensitivity analyses were undertaken (PFS2 and PFS3) that included defining additional clinical outcomes as events and evaluated the impact of informative censoring, an overview of which is shown in Table 14.

Table 13. Summary of the statistical analyses undertaken in the CELESTIAL trial

Hypothesis objective	Statistical analysis	Sample size, power calculation
The null hypothesis was that there was no difference in the duration of OS between the treatment groups (cabozantinib plus BSC versus placebo plus BSC) The alternative hypothesis was that there was a difference in the duration of OS between the treatment groups (cabozantinib plus BSC versus placebo plus BSC)	**Primary efficacy analyses** *Primary efficacy endpoint: OS* Analyses: Up to three analyses were planned: two interim analyses and a final analysis when approximately 50%, 75% and 100% of the total required number of deaths, respectively, were observed, i.e., 311, 466 and 621 deaths, respectively. Hypothesis testing was performed using the stratified log-rank test with a two-sided α=0.05. The stratification factors were the same as those used to stratify randomisation (IxRS data were used). Median duration of OS and the associated 95% confidence interval (CI) for each treatment group was estimated using the Kaplan-Meier method. The stratified HR and its 95% CI were estimated using a Cox proportional hazard model with treatment group as the independent variable and stratified by the randomisation/log-rank test stratification factors. Inflation of Type I error associated with interim analyses was controlled using a Lan-DeMets O'Brien-Fleming alpha-spending function. The calculated critical p-values (and observed hazard ratios [HR]) for rejecting the null hypothesis were 0.0031 (HR ≤0.70), 0.0183 (HR ≤0.80) and 0.044 (HR ≤0.84) for 311, 466 and 621 deaths (50%, 75% and 100% of deaths), respectively. The actual critical values were based on the actual number of events observed at the time of each analysis. The actual critical value for the first interim analysis was 0.0037 (321 deaths, 52% of the total required number of deaths) and for the second interim analysis was 0.021 (484 deaths, 78% of the total required number of deaths). 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 OS 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 OS was rejected in favour of cabozantinib. Formal futility analyses were not planned. *Secondary efficacy endpoint* PFS: investigator-determined radiographical progression according to RECIST 1.1 (only adequate tumour assessments [ATAs] were considered) or death.	The sample size was based on the primary efficacy endpoint (OS). A sample size of 760 patients and 621 events would provide 90% power for a two-sided log-rank test at 5% significance to detect a 31.6% increase in OS with cabozantinib compared with placebo (HR 0.76). Assuming a median OS of 8.2 months in the placebo group (based upon the placebo-controlled brivanib BRISK trial in patients who were previously treated with sorafenib [48]) and exponential distribution, this would correspond to median OS of 10.8 months in the cabozantinib group. The minimum observed effect that would result in statistical significance for OS at the two interim analyses and the final analysis were 42.1% improvement (HR 0.70, i.e. from 8.2 to 11.7 months), 25.7% improvement (HR 0.80, i.e. from 8.2 to 10.3 months) and 18.4% improvement (HR 0.84, i.e. from 8.2 to 9.7 months), respectively.

Hypothesis objective	Statistical analysis	Sample size, power calculation
	ORR: the proportion of patients with a CR or PR as the investigator-determined BOR in terms of tumour assessment category (CR, PR, stable disease, progressive disease or not evaluable) according to RECIST 1.1 that occurred prior to any censoring relevant for the primary analysis of PFS (see Table 14 for censoring details). Analysis of the secondary endpoints only took place if the result of either an interim analysis or the final analysis of OS achieved statistical significance compared with placebo. The hypotheses for PFS and ORR were tested in parallel; PFS was tested with a two-sided α=0.04 and ORR with a two-sided α=0.01. The primary analysis of PFS was performed in a similar manner to the primary analysis of OS For BOR, confirmation of response was required ≥28 days after the response was first observed. Hypothesis testing for ORR was performed using Fisher exact test. Analysis using the Cochran-Mantel-Haenszel (CMH) method to adjust for randomisation stratification factors was also performed. If the ORR was >10%, the duration of the objective response and time to the objective response were calculated. The duration of objective response (the time from the first documentation of objective response by the investigator, confirmed ≥28 days later, to disease progression or death due to any cause) was calculated using the Kaplan-Meier method with the dates of progression and censoring determined as described for the analysis of PFS. The time to objective response was the time from randomisation to the first documentation of objective response by the investigator, which was confirmed ≥28 days later. *Multiplicity* The multiplicity issue resulting from analysis of one primary endpoint, two secondary efficacy endpoints (PFS and ORR) and planning two interim analyses for testing OS was addressed by employing a fixed-sequence testing procedure, applying a modified Bonferroni procedure (dividing the α between the secondary endpoints), and implementing an α-spending function. *Exploratory endpoints* Safety was analysed descriptively. In general, other than for partial dates, missing data were not imputed.	

Abbreviations: BOR, best overall response; BSC, best supportive care; CR, complete response; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PR, partial response.
Source: Exelixis, 2018 [46].

Table 14. Event and censoring rules for the primary analysis of PFS (PFS1) and the sensitivity analyses (PFS2 and PFS3)

Analysis	PFS1		PFS2		PFS3	
Purpose	Primary		Sensitivity		Sensitivity	
Situation	Outcome	Date	Outcome	Date	Outcome	Date
No post-baseline assessment	Censored	Date of randomisation	Censored	Date of randomisation	Censored	Date of randomisation
Radiographical PD	Event	Date of PD	Event	Date of PD	Event	Date of PD
Death	Event	Date of PD	Event	Date of PD	Event	Date of PD
Subsequent systemic or local liver-directed NPACT	Censored	Date of last ATA on or prior to date of NPACT	Event	Date of NPACT	Censored	Date of last ATA on or prior to date of NPACT
Radiation (other than to bone)	Censored	Date of last ATA on or prior to date of radiation	Event	Date of radiation	Censored	Date of last ATA on or prior to date of radiation
Surgery to resect tumour lesions	Censored	Date of last ATA on or prior to date of surgery	Event	Date of surgery	Censored	Date of last ATA on or prior to date of surgery
Event after >2 missed ATAs (>126 days)	Censored	Date of last ATA prior to the missing visits	Censored	Date of last ATA prior to the missing visits	Censored	Date of last ATA prior to the missing visits
Treatment discontinuation due to clinical deterioration	NA	NA	Event	Date of determination	Event	Date of determination
No event by last ATA	Censored	Date of last ATA	Censored	Date of last ATA	Censored	Date of last ATA

Abbreviations: ATA, adequate tumour assessments; NPACT, non-protocol anticancer therapy; PD, progressive diseases; PFS, progression-free survival.
Source: Exelixis, 2018 [46].

B.3.4.3 Participant flow in the CELESTIAL trial

See the CONSORT diagram for the CELESTIAL trial in Appendix D [7].

B.3.5 Quality assessment of the relevant clinical effectiveness evidence

A quality assessment of the CELESTIAL trial is summarised in Table 15. The CELESTIAL 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 15. Quality assessment results for the CELESTIAL trial

Trial	The CELESTIAL 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.

B.3.6 Clinical effectiveness results of the relevant trials

B.3.6.1 Primary endpoint: OS

The data presented are from the second interim analysis, planned for when 75% of the total number of required deaths to adequately power the trial (621 deaths), i.e., 466 deaths, had occurred [7, 46]. At the cut-off date for the second interim analysis (1 June 2017), 484 deaths in the overall population had been reported, representing 78% of the total number of deaths required. The median duration of follow-up for OS was 22.9 months. Cabozantinib significantly reduced the risk of death by 24% compared with placebo (HR 0.76 [95% CI: 0.63, 0.92]; stratified log-rank p-value 0.005) increasing the median OS by 2.2 months (10.2 versus 8.0 months) (Table 16; Figure 4). The landmark estimate of the proportion of patients alive at 12 months was 46% in the cabozantinib group compared with 34% in the placebo group [7].

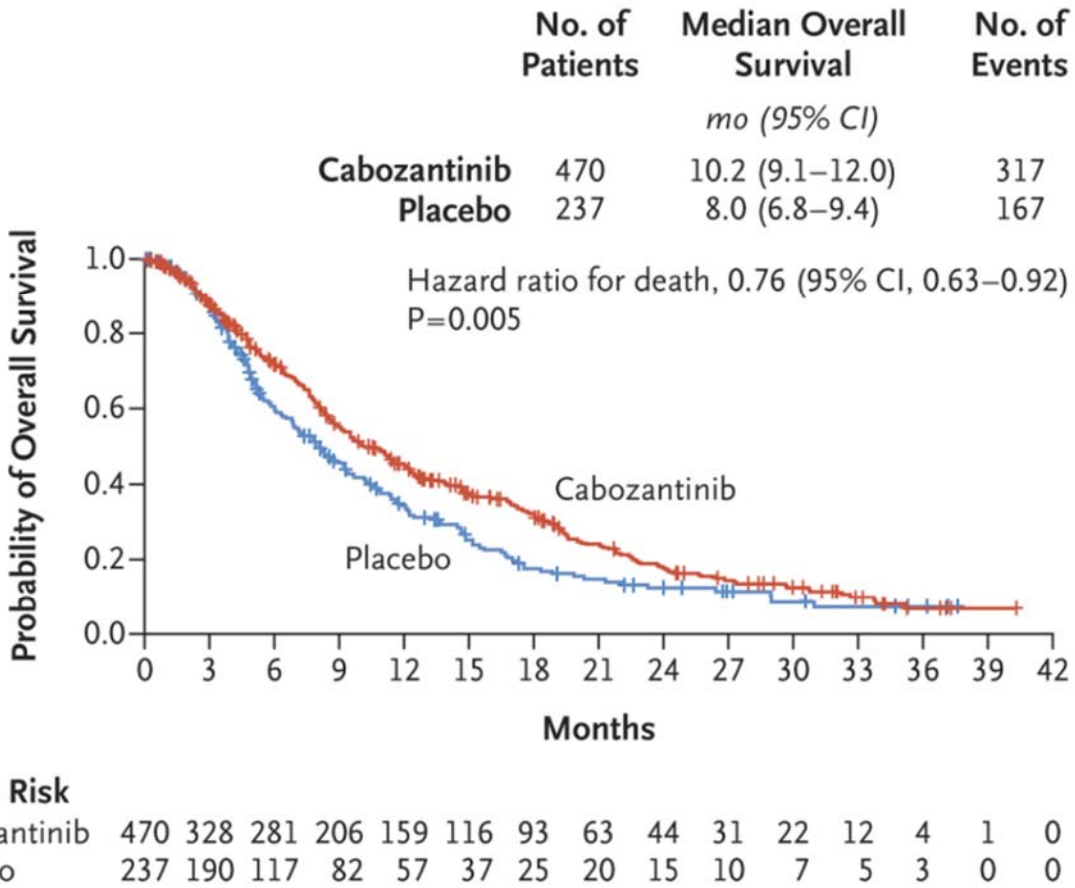
Thus, the null hypothesis that there was no difference in the duration of OS between the treatment groups (cabozantinib plus BSC versus placebo plus BSC) was rejected as a result of the second interim analysis. As a result of this no further analyses of OS were planned.

Table 16. The CELESTIAL trial: duration of OS (ITT; second planned interim analysis)

	Cabozantinib (n=470)	Placebo (n=237)
Patients, n (%)		
Censored	153 (33)	70 (30)
Death	317 (67)	167 (70)
Duration of OS (months)		
Median (95% CI)	10.2 (9.1, 12.0)	8.0 (6.8, 9.4)
Range	0.1, 40.3+	0.03+, 37.6+
Critical p-value to reject null hypothesis of equal OS	0.02	
Observed p-value (stratified log-rank test)	0.005	
HR (95% CI; stratified)	0.76 (0.63, 0.92)	
Observed p-value (unstratified log-rank test)	0.0072	
HR (95% CI; unstratified)	0.77 (0.64, 0.93)	

+ indicates a censored observation
Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; OS, overall survival.
Source: Abou-Alfa et al., 2018 [7]. Exelixis, 2018 [46].

Figure 4. The CELESTIAL trial: OS with cabozantinib versus placebo – Kaplan-Meier plot (ITT population; second planned interim analysis, adjusted)



Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; OS, overall survival.
Source: Abou-Alfa et al., 2018 [7].

B.3.6.2 Secondary endpoint: PFS

Analysis of PFS was conducted in the ITT population at the time of the primary analysis of the primary endpoint OS, i.e., at the time of the second planned interim analysis, due to the significant result for the primary endpoint [7, 46]. In the pre-specified primary analysis of the secondary efficacy endpoint PFS, PFS was defined as the time from randomisation to investigator-determined radiographical progression according to RECIST 1.1 or death due to any cause in the ITT population.

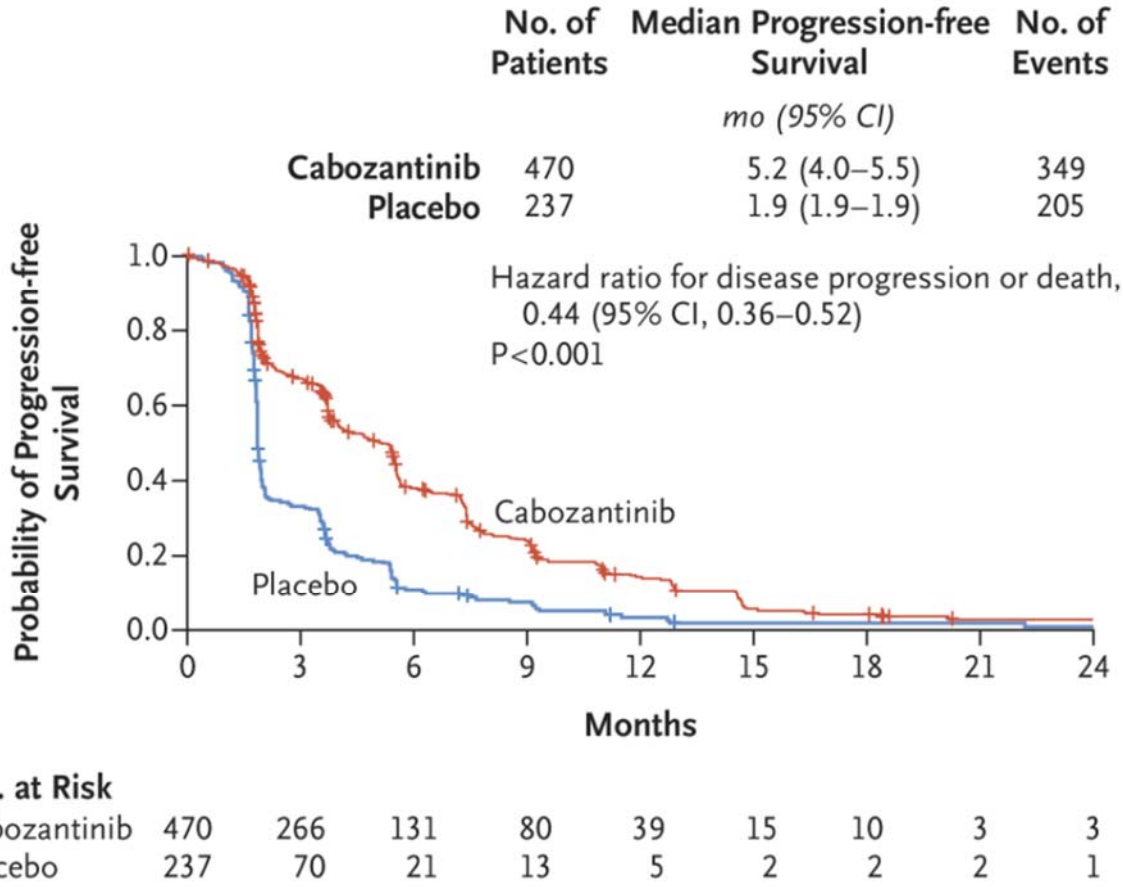
Cabozantinib significantly reduced the risk of disease progression/death by 56% compared with placebo (HR 0.44 [95% CI: 0.36, 0.52]; stratified log-rank p-value <0.0001) increasing median PFS by 3.3 months (5.2 versus 1.9 months) at the time of the second planned interim analysis (Table 17; Figure 5). The landmark estimate of the proportion of patients alive and progression-free at 12 months was 15% in the cabozantinib group compared with 3% in the placebo group.

Table 17. The CELESTIAL trial: PFS (investigator assessed; ITT population; second interim analysis)

	Cabozantinib (n=470)	Placebo (n=237)
Number (%) of patients		
Censored	121 (26)	32 (14)
Event	349 (74)	205 (86)
Death	65 (14)	19 (8.0)
PD	284 (60)	186 (78)
Duration of PFS (months)		
Median (95% CI)	5.2 (4.0, 5.5)	1.9 (1.9, 1.9)
Range	0.03+, 33.2	0.03+, 25.5+
Critical p-value to reject null hypothesis of equal PFS	0.04	
Observed p-value (stratified log-rank test)	<0.0001	
HR (95% CI; stratified)	0.44 (0.36, 0.52)	
Observed p-value (unstratified log-rank test)	<0.0001	
HR (95% CI; unstratified)	0.46 (0.38, 0.55)	

+ indicates a censored observation
Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; PD, progressive disease; PFS, progression-free survival.
Source: Abou-Alfa et al., 2018 [7]. Exelixis, 2018 [46].

Figure 5. The CELESTIAL trial: PFS – Kaplan-Meier plot (investigator assessed; ITT population; second interim analysis, adjusted)



Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; PFS, progression-free survival.
Source: Abou-Alfa et al., 2018 [7].

The robustness of the significant improvement in PFS with cabozantinib compared with placebo was confirmed in the unadjusted analysis and in sensitivity analyses. The results of two sensitivity analyses (PFS2 and PFS3; data not used in the economic model) in which PFS was defined using additional clinical outcomes as events and which also evaluated the impact of informative censoring were similar to those in the primary analysis (Table 18).

Table 18. The CELESTIAL trial: results of sensitivity analyses for PFS (investigator assessed; ITT population; second interim analysis)

PFS analysis	Cabozantinib (n=470)		Placebo (n=237)		Cabozantinib versus placebo	
	Events, % (n)	Mean, months	Events, % (n)	Mean, months	HR (95% CI) stratified	p-value log-rank test, stratified
Primary analysis	74 (349)	5.2	86 (205)	1.9	0.44 (0.36, 0.52)	<0.0001
Sensitivity analyses						

PFS analysis	Cabozantinib (n=470)		Placebo (n=237)		Cabozantinib versus placebo	
	Events, % (n)	Mean, months	Events, % (n)	Mean, months	HR (95% CI) stratified	p-value log-rank test, stratified
PFS2	80 (374)	4.4	89 (211)	1.9	0.46 (0.38, 0.55)	<0.0001
PFS3	76 (356)	4.7	87 (207)	1.9	0.44 (0.37, 0.53)	<0.0001

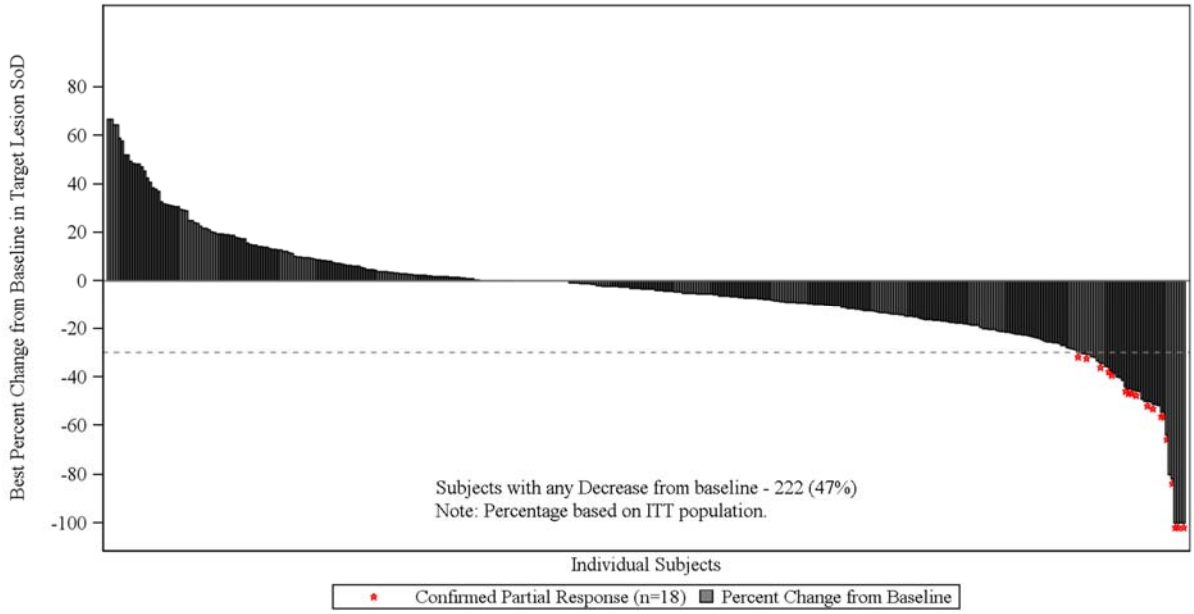
Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; OS, overall survival; PFS, progression-free survival.
Source: Abou-Alfa et al., 2018 [7] Exelixis, 2018 [46]

B.3.6.3 Secondary endpoint: ORR

Analysis of the secondary efficacy endpoint ORR (investigator-determined CR or PR according to RECIST 1.1) was conducted in the ITT population at the time of the primary analysis of OS, i.e., at the time of the second planned interim analysis, due to the significant result for the OS [7, 46].

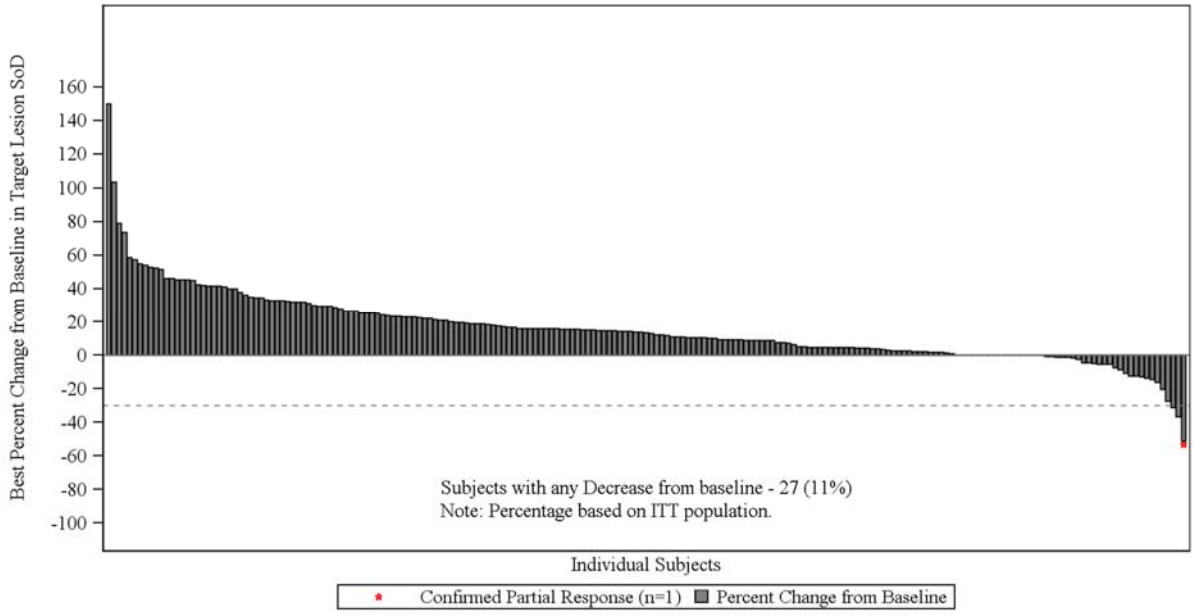
The best percentage change from baseline in tumour target lesion size (investigator-determined according to RECIST 1.1) is depicted in Figure 6 (cabozantinib) and Figure 7 (placebo). Post-baseline reduction in the sum of target lesion diameters (SoD) was observed in 47% of subjects in the cabozantinib arm and 11% in the placebo arm. The waterfall plots do not include subjects which lack of evaluable post-baseline assessment, censoring (per PFS rules) before first evaluable post-baseline assessment, lack of target lesions, and/or incomplete or unevaluable target lesion assessment. Data from time points after the first date of any of the censoring events defined for the primary PFS analysis were also excluded from the plots.

Figure 6. Waterfall plot of best percentage change in tumour target lesion size from baseline per Investigator; Cabozantinib arm (ITT population, subjects with a baseline and post-baseline target lesion assessment, N = 388)



Abbreviations: ITT, intention to treat; SoD, sum of target lesion diameters.
Source: Exelixis, 2018 [46].

Figure 7. Waterfall plot of best percentage change in tumour target lesion size from baseline per investigator; Placebo arm (ITT population, subjects with a baseline and post-baseline target lesion assessment, N = 205)



Abbreviations: ITT, intention to treat; SoD, sum of target lesion diameters.
Source: Exelixis, 2018 [46].

The results for BOR clearly demonstrate a higher disease control rate with cabozantinib compared with placebo (64% versus 33%). Cabozantinib was associated with a significantly higher ORR than placebo (odds ratio [OR] 9.4 [95% CI 1.2, 71.0]; stratified Cochran-Mantel-Haenszel (CMH) p-value 0.0086). As no patient in either treatment group had a CR, these results reflect the significantly higher PR rate with cabozantinib compared with placebo (4% versus 0.4%). As would be expected due to the significantly higher ORR with cabozantinib, cabozantinib was also associated with a lower rate of progressive disease (PD) compared with placebo (21% versus 55%) (Table 19).

Table 19. The CELESTIAL trial: ORR for cabozantinib versus placebo (investigator-determined; ITT population; second interim analysis)

	Cabozantinib n=470	**Placebo n=237**
BOR, n (%)		
Confirmed CR	0	0
Confirmed PR	18 (4)	1 (0.4)
SD	282 (60)	78 (33)
Unconfirmed CR	0	0
Unconfirmed PR	13 (3)	2 (0.8)
PD	98 (21)	131 (55)
Unable to evaluate/missing	72 (15)	27 (11)
No baseline assessment	0	0
No post-baseline assessments	65 (14)	22 (9)
No qualifying post-baseline assessment on or before primary PFS analysis censoring or event	7 (1)	5 (2)
ORR [CR + PR], n (%)	18 (4)	1 (0.4)
95% CI	(2.3, 6.0)	(0.0, 2.3)
Treatment difference (cabozantinib – placebo) (95% CI)	3.4 (1.49, 5.33)	
Critical p-value to reject null hypothesis of equal ORR	0.01	
Observed p-value (stratified CMH test)	0.0086	
Odds ratio, stratified (95% CI)	9.4 (1.2, 71.0)	
Observed p-value (unstratified Fishers exact test)	0.0059	
Odds ratio, unstratified (95% CI)	9.4 (1.2, 70.8)	

Abbreviations: BOR, best overall response; CI, Confidence Interval; CMH, Cochran-Mantel-Haenszel; CR, complete response; ITT, intention to treat; ORR, objective response rate; OS, overall survival; PD, progressive disease; PFS, progression-free survival; PR, partial response; SD, stable disease
Source: . Abou-Alfa et al., 2018 [7], Exelixis, 2018 [46]

B.3.6.4 Exploratory endpoint: safety, including TTD

In the CELESTIAL trial, patients received cabozantinib for almost twice as long as patients received placebo: the median duration of exposure at the time of the planned second interim analysis of OS (cut-off date 1 June 2017) was 3.8 months (range 0.1, 37.3) in the cabozantinib group compared with 2.0 months (range 0.0, 27.2) in the placebo group [7, 46].

Data regarding AEs are reported in Section B.3.8.

B.3.7 Subgroup analysis

There are no subgroups of interest as the target population is the full marketing authorisation. In an ad hoc subgroup analysis, subjects whose only prior therapy for HCC was sorafenib also showed an OS benefit. Subgroup analyses demonstrated a generally consistent OS and PFS benefit for cabozantinib treated patients in all subgroups comprising at least 20 patients. There were too few responders to interpret ORR subgroup analyses. The CELESTIAL study was not powered to assess differential patient response to treatment in subgroups.

More detailed results of the subgroup analysis are provided in Appendix E.

B.3.8 Adverse reactions

B.3.8.1 Summary of safety data

In the CELESTIAL trial, the population for the analysis of safety (safety population) comprised of all patients who received at least one dose of study drug (n=704; n=467 for cabozantinib and n=237 for placebo). In the safety population in the CELESTIAL trial, patients in the placebo group received a mean (±standard deviation) daily dose of 52.85 mg (±11.1) and those in the cabozantinib group received a mean daily dose of 36.56 mg (±13.8) [46].

Cabozantinib was generally well tolerated. AEs frequently reported with cabozantinib were typical of those with VEGFR-TKI therapies. An overview of safety data from the CELESTIAL trial is provided in Table 20 and Table 21.

Table 20: The CELESTIAL trial: summary of safety data (safety population)

Adverse Events	Cabozantinib n=467 n (%)	Placebo n=237 n (%)
Any AE (all grades)	460 (99)	219 (92)
Grade 3 or 4 AEs	316 (68)	86 (36)
Treatment-related AEs	439 (94)	148 (62)
SAEs	232 (50)	87 (37)
Treatment-related SAEs	82 (18)	14 (5.9)
Treatment-related Grade 5 AE	6 (1.3)	1 (0.4)
Deaths (at any time, excluding PD)	314 (67)	167 (70)
AE leading to dose modification	416 (89)	94 (40)
AE leading to discontinuation of study drug	96 (21)	10 (4.2)

Abbreviations: AEs, adverse events; PD, progressive disease; SAEs, serious adverse events.
Source: Abou-Alfa et al., 2018 [7], Exelixis, 2018 [46]

Table 21. AEs* (any grade) reported in ≥10% of patients in either treatment group

Event	Cabozantinib (number of patients (percent)			Placebo (number of patients (percent)		
	Any grade	Grade 3	Grade 4	Any grade	Grade 3	Grade 4
Any AE	460 (99)	270 (58)	46 (10)	219 (92)	80 (34)	6 (3)
Diarrhoea	251 (54)	45 (10)	1 (<1)	44 (19)	4 (2)	0
Decreased appetite	225 (48)	27 (6)	0	43 (18)	1 (<1)	0
PPES	217 (46)	79 (17)	0	12 (5)	0	0
Fatigue	212 (45)	49 (10)	0	70 (30)	10 (4)	0
Nausea	147 (31)	10 (2)	0	42 (18)	4 (2)	0
Hypertension	137 (29)	73 (16)	1 (<1)	14 (6)	4 (2)	0
Vomiting	121 (26)	2 (<1)	0	28 (12)	6 (3)	0
Increase in AST level	105 (22)	51 (11)	4 (1)	27 (11)	15 (6)	1 (<1)
Asthenia	102 (22)	31 (7)	1 (<1)	18 (8)	4 (2)	0
Dysphonia	90 (19)	3 (1)	0	5 (2)	0	0
Constipation	87 (19)	2 (<1)	0	45 (19)	0	0
Abdominal pain	83 (18)	7 (1)	1 (<1)	60 (25)	10 (4)	0
Weight loss	81 (17)	5 (1)	0	14 (6)	0	0
Increase in ALT level	80 (17)	23 (5)	0	13 (5)	5 (2)	0
Mucosal inflammation	65 (14)	8 (2)	0	5 (2)	1 (<1)	0
Pyrexia	64 (14)	0	0	24 (10)	1 (<1)	0
Upper abdominal pain	63 (13)	3 (1)	0	31 (13)	0	0
Cough	63 (13)	1 (<1)	0	26 (11)	0	0
Peripheral oedema	63 (13)	4 (1)	0	32 (14)	2 (1)	0
Stomatitis	63 (13)	8 (2)	0	5 (2)	0	0
Dyspnoea	58 (12)	15 (3)	0	24 (10)	1 (<1)	0
Rash	58 (12)	2 (<1)	0	14 (6)	1 (<1)	0
Ascites	57 (12)	17 (4)	1 (<1)	30 (13)	11 (5)	0
Dysgeusia	56 (12)	0	0	5 (2)	0	0
Hypoalbuminemia	55 (12)	2 (<1)	0	12 (5)	0	0
Headache	52 (11)	1 (<1)	0	16 (7)	1 (<1)	0
Headache	52 (11)	1 (<1)	0	16 (7)	1 (<1)	0
Insomnia	49 (10)	1 (<1)	0	17 (7)	0	0
Dizziness	48 (10)	2 (<1)	0	15 (6)	0	0
Dyspepsia	47 (10)	0	0	7 (3)	0	0
Anaemia	46 (10)	18 (4)	1 (<1)	19 (8)	12 (5)	0
Back pain	46 (10)	5 (1)	0	24 (10)	1 (<1)	0
Increase in serum bilirubin level	45 (10)	10 (2)	4 (1)	17 (7)	2 (1)	2 (1)
Decrease in platelet count	45 (10)	13 (3)	4 (1)	7 (3)	2 (1)	0

* Listed are adverse events, regardless of causality, that were reported in at least 10% of patients in either group. Severity was graded according to National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0.

Abbreviations: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; PPES, palmar-plantar erythrodysaesthesia syndrome.

Source: Abou-Alfa et al., 2018 [7].

The rate of discontinuation of cabozantinib or placebo owing to adverse events that were considered to be related to the trial regimen was 16% (76 patients) in the cabozantinib group and 3% (7 patients) in the placebo group. Adverse events leading to treatment discontinuation in more than 1.0% of patients in the cabozantinib group were palmar-plantar erythrodysesthesia, fatigue, decreased appetite, diarrhoea, and nausea [7].

AEs of any grade regardless of causality were reported in 99% of the patients in the cabozantinib group and in 92% in the placebo group, and AEs of grade 3 or 4 were reported in 68% of the patients in the cabozantinib group and in 36% in the placebo group (Table 21). The most common grade 3 or 4 AEs in the cabozantinib group were palmar-plantar erythrodysesthesia (17%, vs. 0% with placebo), hypertension (16% vs. 2%), increased aspartate aminotransferase level (12% vs. 7%), fatigue (10% vs. 4%), and diarrhoea (10% vs. 2%). The most common AEs of any grade leading to dose reductions in the cabozantinib group were palmar-plantar erythrodysesthesia (22%), diarrhoea (10%), fatigue (7%), hypertension (7%), and increased aspartate aminotransferase level (6%). Serious AEs were reported in 50% of the patients who received cabozantinib and in 37% of the patients who received placebo. A serious AE was defined as an AE of any grade that caused death, was life-threatening, resulted in hospitalisation or prolongation of hospitalisation, was deemed medically important, or resulted in disability or birth defect. Grade 5 AEs occurring within 30 days after the last dose of cabozantinib or placebo were reported in 55 patients (12%) in the cabozantinib group and in 28 patients (12%) in the placebo group and were commonly related to disease progression [7].

Grade 5 AEs that were considered to be related to cabozantinib or placebo were reported in 6 patients in the cabozantinib group (one event each of hepatic failure, tracheoesophageal fistula, portal-vein thrombosis, upper gastrointestinal haemorrhage, pulmonary embolism, and the hepatorenal syndrome) and in 1 patient in the placebo group (hepatic failure) [7].

B.3.8.4 Overview of the safety of the technology in relation to the decision problem

Cabozantinib has been licensed and marketed in the US since 2016, in Europe since 2016 for renal cell carcinoma and for HCC since November 2018. AEs in patients participating in the CELESTIAL trial were as expected in those with pre-treated advanced HCC. AEs characteristic of HCC in the context of chronic liver disease/cirrhosis were observed with cabozantinib and placebo and Grade 3 and 4 AEs associated with advanced HCC or underlying liver disease were reported frequently.

It is anticipated that cabozantinib will have an acceptable, recognisable, and manageable safety profile when used in the context of the decision problem.

Further details of AEs reported in the CELESTIAL study are provided in Appendix F.

B.3.9 Meta-analysis

No meta-analysis was carried out, as the only two trials identified as relevant to the decision problem were the CELESTIAL trial that compared cabozantinib with placebo, and the RESORCE trial that compared regorafenib with placebo.

B.3.10 Indirect and mixed treatment comparisons

For reasons detailed in Section B.2, regorafenib has been selected as the reference comparator for the cost-comparison analysis. The FTA framework suggests that the technology should have similar efficacy to the comparator. In the absence of a head-to-head trial comparing cabozantinib with regorafenib, the following ITCs have been conducted to estimate the relative efficacy of cabozantinib versus regorafenib:

- One based on Bucher et al. [49], and
- The other being a matching-adjusted indirect comparison (MAIC).

The CELESTIAL and RESORCE trials were identified as the only relevant trials to perform the indirect comparisons and both trials shared a common comparator treatment, placebo. The summary of these trials is included in Table 22.

B.3.10.1 Identification of studies

The systematic literature review (SLR) described in Appendix D, was used to identify all potential studies that may have been relevant for indirect comparison with cabozantinib.

Table 22. Summary of the trials used to carry out the indirect treatment comparison

Trial reference	CELESTIAL	RESORCE
Intervention (N)	Cabozantinib (60 mg qd) plus BSC (470)	Regorafenib (160 mg qd) plus BSC (379) - once daily during weeks 1–3 of each 4-week cycle
Comparator (N)	Placebo plus BSC (237)	Placebo plus BSC (194) - once daily during weeks 1–3 of each 4-week cycle
Study initiation and completion (years)	26 September 2013 – 01 June 2017 (data cut-off date)	May 2013 – Feb 2016 (primary completion date)
Phase	III	III
Patient population (ITT)	Sorafenib tolerant and intolerant; second and third-line patients (CELESTIAL inclusion criteria listed in Table 9)	Sorafenib tolerant, second-line patients only
Method of blinding	Double-blind	Double-blind
Randomisation	2:1, stratified by etiology of disease (HBV [with or without HCV], HCV [without HBV], or Other), geographic region (Asia, Other Regions), and presence of extrahepatic spread of disease and/or macrovascular invasion (yes versus no).	2:1, stratified by geographical region (Asia versus rest of world), macrovascular invasion (yes versus no), extrahepatic disease (yes versus no), α-fetoprotein concentration (<400 ng/mL versus ≥400 ng/mL), and ECOG performance status (0 versus 1).
Study centres	Multicentre (Europe, North America, Australia, New Zealand, Asia)	Multicentre (Europe, North America, Australia, South America, Asia)
Median follow-up duration	22.9 months	7.0 months
Patients censored for OS (%)	32%	37%

Abbreviations: BSC, best supportive care; ECOG, Eastern Cooperative Oncology Group; ITT, intention to treat; N, number of participants; OS, overall survival; qd, once a day.
Source: Abou-Alfa et al., 2018 [7], Exelixis, 2018 [46], Finn et al, 2018 [50]

B.3.10.2 Indirect treatment comparison based on Bucher et al methodology

An ITC based on the approach used by Bucher et al. [49] was performed to estimate the relative efficacy of cabozantinib versus regorafenib, in accordance with the decision problem outlined in Table 1. The principal assumption of the Bucher ITC is that the relative efficacy of the treatments included in the comparison is the same in all trials included in the indirect comparison. To satisfy this assumption, the trials need to be comparable in terms of study design and patient characteristics. For this

analysis, however, it should be noted that the Bucher approach is limited by the fact that the ITT population results for the overall population of CELESTIAL trial which included second and third-line patients would be compared against the overall population of the RESORCE trial which includes second-line patients only.

Comparison of trial design and patient characteristics

Both trials (CELESTIAL and RESORCE) were phase III, multicentre, double-blind RCTs, conducted over similar durations and in similar geographical locations suggestive of consistent clinical practices across both trials. However, the trials populations differed in several baseline characteristics with differences in the ethnic mix, region, ECOG performance status, number of prior treatments and duration of prior sorafenib treatment between the trial populations. A comparison of baseline characteristics showed that, on average, patients enrolled in the CELESTIAL trial had a shorter duration of prior sorafenib treatment than patients in RESORCE (8 versus 12 months). Additionally, patients in CELESTIAL were less likely to have an ECOG PS of 0 (53% versus 66%), more likely to be white (56% versus 36%), and less likely to be in the Asia geographical region (25% versus 38%). The baseline characteristics from CELESTIAL and RESORCE trial are presented in Table 23.

Table 23. Comparison of baseline characteristics of subjects enrolled in CELESTIAL and RESORCE

	CELESTIAL	RESORCE
Treatment (N)	**Cabozantinib (N = 470)**	**Regorafenib (N = 374)**
Age under 65	51	55
Female	18	12
Asia geographical region	25	38
ECOG status 0	53	66
Child-Pugh class A	100	98
Mean duration of sorafenib treatment (months)	8	12
Extrahepatic disease	78	72
Macrovascular invasion	30	29
Hepatitis B aetiology	38	38
Alcohol use aetiology	22	25
Hepatitis C aetiology	24	21
AFP > 400ng/mL	41	43
White (%)	56	36

Abbreviations: AFP; alpha fetoprotein; ECOG, Eastern Cooperative Oncology Group; N, number of patients.
Source: Abou-Alfa et al., 2018 [7], Exelixis, 2018 [46], Finn et al, 2018 [50].

Methodology

The effect of cabozantinib relative to regorafenib was estimated using the method for adjusted indirect comparison developed by Bucher et al. [49]. The method applies aggregate data from the CELESTIAL and RESORCE trials, with the placebo plus BSC as the common comparator arm, to derive the indirect estimators of the efficacy of cabozantinib relative to regorafenib for the outcomes of interest. The method allows the randomisation of the RCTs to be preserved by utilising the relative treatment effects from each of the randomised trials. The main underlying assumption is that there is no difference in the distribution of effect modifying variables between trials, which allows the combination of their relative effects. The Bucher ITC for cabozantinib versus regorafenib included the OS primary, PFS secondary endpoints and safety of both trials [49].

Results – efficacy outcomes

The results of the Bucher ITC showed hazard ratios versus regorafenib that favoured cabozantinib for PFS [HR 0.96 (0.73, 1.26)] and that favoured regorafenib for OS [HR 1.21 (0.90, 1.62)], but the results were not statistically significant suggesting similar efficacy in terms of OS and PFS for both treatments.

The efficacy results are presented in Table 24.

Table 24: Summary of Bucher ITC results for cabozantinib plus BSC versus regorafenib plus BSC in ITT populations of their respective trials

Endpoint: relative effect measure	CELESTIAL: Cabozantinib versus placebo HR (95% CI)	RESORCE: Regorafenib versus placebo HR (95% CI)	Bucher ITC: Cabozantinib versus regorafenib HR (95% CI)
Overall survival	0.76 (0.63, 0.92)	0.63 (0.50, 0.79)	1.21 (0.90, 1.62)
Progression-free survival	0.44 (0.36, 0.52)	0.46 (0.37, 0.56)	0.96 (0.73, 1.26)

Abbreviations: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison.
Source: Abou-Alfa et al., 2018 [7], Exelixis, 2018 [46], Finn et al, 2018 [50]

Log cumulative hazard plots and Schoenfeld residual plots were used to test the proportional hazards assumption underlying the Bucher ITC. Therefore, OS and PFS Kaplan-Meier curves from RESORCE were digitised and pseudo individual patient level data (IPD) generated, using the Guyot algorithm [51]. The curves in the log cumulative hazard plot for OS were not parallel and cross (Figure 8). Furthermore, the Schoenfeld residuals show correlation with time (Figure 9) and a Grambsch and Therneau test (a more formal statistical test based on the scaled Schoenfeld residuals)

had a p-value of 0.0016. The findings suggested that the proportional hazards assumption was not satisfied for OS.

Figure 8: OS Log-log cumulative hazards

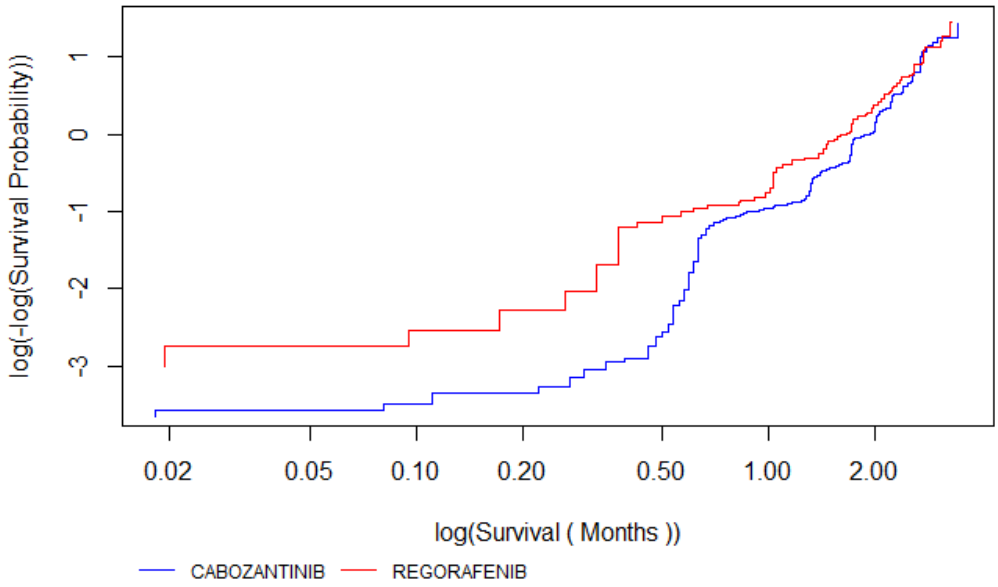
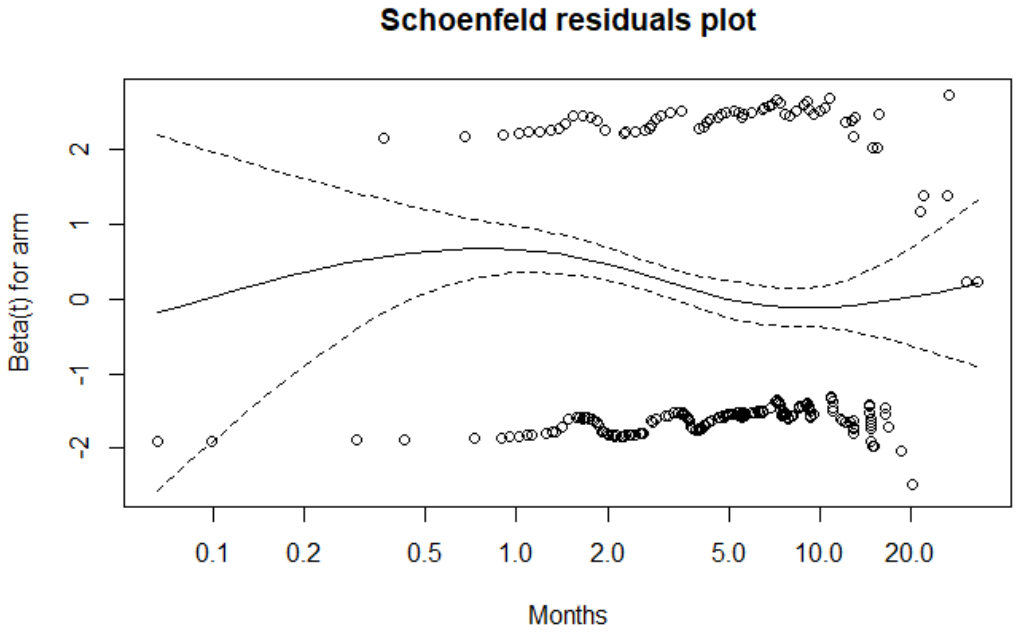


Figure 9: OS Schoenfeld residuals plot



However, for PFS, the curves in the log cumulative hazard plot were overlapping (Figure 10). The Schoenfeld residuals showed little correlation with time (Figure 11) and the Grambsch and Therneau test shows a p-value of 0.73. The findings suggested that the proportional hazards assumption was satisfied for PFS.

Figure 10: PFS Log-log cumulative hazards

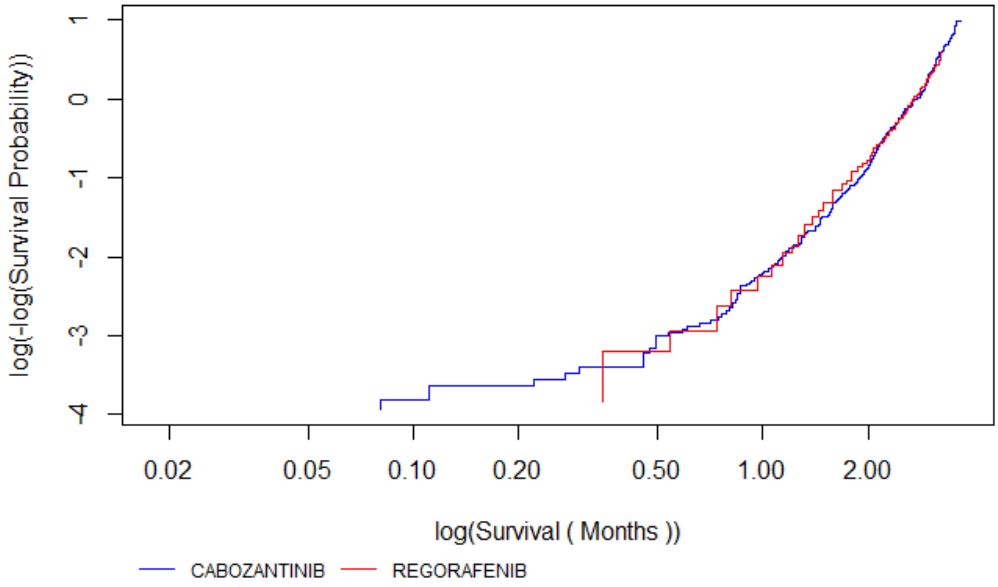
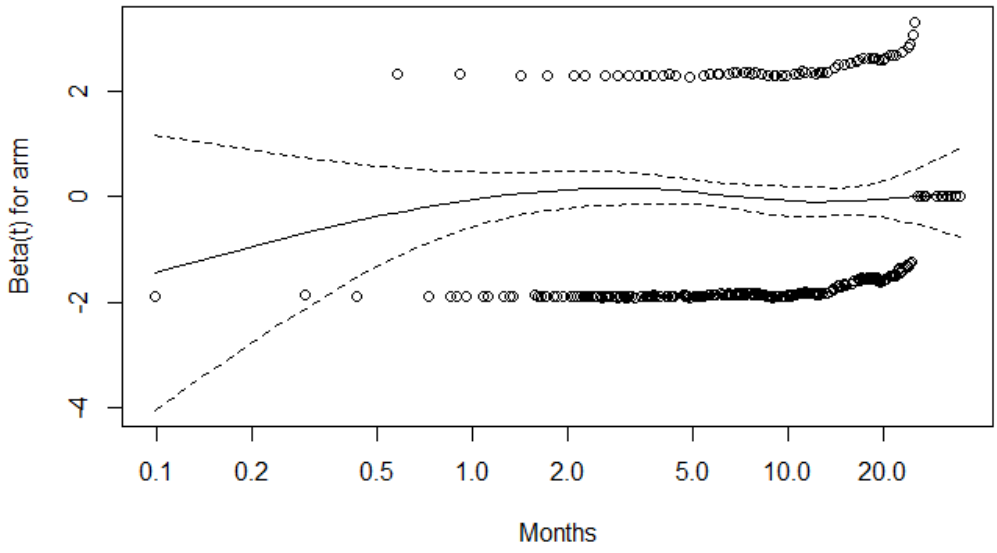


Figure 11: PFS Schoenfeld residuals plot



Results – safety outcomes

Treatment-emergent AEs with a grade 3/4 that occurred in ≥5% of patients in either arm was analysed. This is considered a standard approach as treatment-emergent grade 3/4 events are likely to be associated with higher costs and larger impact on quality of life than grade 1/2 events. This is consistent with previous submission to NICE in advanced HCC [2]. Table 25 below presents the AEs considered in the analysis.

Table 25: Treatment-emergent AEs with a grade 3/4 that occurred in ≥5% of patients from CELESTIAL and RESORCE

Adverse Events	Cabozantinib n=467 n (%)	CELESTIAL placebo n=237 n (%)	Regorafenib n=374 n (%)	RESORCE placebo n=193 n (%)
Palmar-plantar erythrodysaesthesia syndrome	78 (16.7)	0 (0)	47 (12.6)	1 (0.5)
Hypertension	69 (14.8)	2 (0.8)	49 (13.1)	6 (3.1)
Elevated aspartate aminotransferase	36 (7.7)	11 (4.6)	19 (5.1)	10 (5.2)
Fatigue	39 (8.4)	6 (2.5)	24 (6.4)	3 (1.6)
Diarrhoea	42 (9.0)	2 (0.8)	9 (2.4)	0 (0)
Elevated bilirubin	0 (0)	0 (0)	25 (6.7)	4 (2.1)

Abbreviations: AEs, adverse events.
Source: Bruix et al,2017 [41], Exelixis, 2018 [46].

For the comparison of AEs, Bucher adjusted comparisons were only feasible when there were events in all arms of CELESTIAL and RESORCE. Therefore, only hypertension, elevated aspartate aminotransferase and fatigue AEs were compared. The results show no statistically significant differences between the AE ORs for cabozantinib and regorafenib. It should be noted that the small number of events results in large confidence intervals. The results are shown in Table 26.

Table 26: Summary of Bucher ITC safety results

Adverse Events	Cabozantinib vs. Regorafenib OR (95% CI)
Hypertension	0.2 (0.0-1.2)
Elevated aspartate aminotransferase	0.6 (0.2-1.6)
Fatigue	1.2 (0.3-5.6)

Abbreviations: CI, confidence interval; ITC, indirect treatment comparison; OR: odds ratio.

B.3.10.3 Matching-adjusted indirect comparison

Given the differences in baseline characteristics between CELESTIAL and RESORCE and the finding that the PH assumption may not be supported for OS, the efficacy of cabozantinib and regorafenib was compared using a MAIC as it provides a method of comparing absolute treatment effects while lowering the risk of bias associated with naïve unadjusted comparisons [52, 53].

The MAIC analysis utilised a subpopulation from the CELESTIAL ITT population, specifically second-line hepatocellular carcinoma patients who had prior treatment

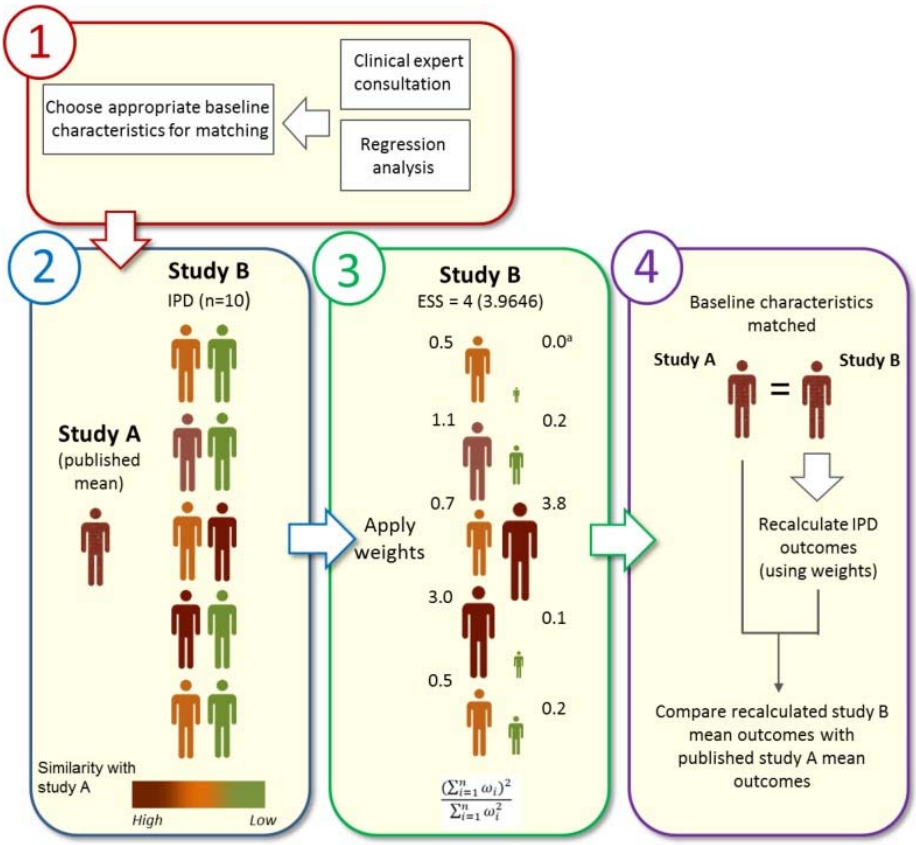
with sorafenib (i.e., pure second-line patients) in order to compare to the RESORCE trial.

The method incorporates IPD, in this case available for CELESTIAL, which were reweighted to mimic the population of the RESORCE trial for which only aggregate results were available. The survival outcomes were recalculated for each pure second-line patient in CELESTIAL using the weighted data.

Methodology

An overview of the MAIC procedure is presented below in Figure 12

Figure 12: Overview of MAIC procedure



Abbreviations: IPD, Individual patient data, ESS: Effective sample size.
Source: Nash et al, 2018 [54].

Baseline characteristics

Comparison of the patient characteristics of RESORCE with those of the pure second-line population of CELESTIAL suggest some differences remained in terms of ethnic mix, region, ECOG performance status and duration of prior sorafenib treatment

between both populations. Other characteristics mentioned in Table 27 were similar or had minor differences.

Two different scenarios were considered to assess the impact of choosing different baseline characteristics for matching:

- In the first scenario (S1), which represents the base case, the baseline characteristics selected for matching were those deemed potential effect modifiers by the clinical experts.
- In the second scenario (S2), which serves as sensitivity check, the baseline characteristics selected for matching were those selected using the stepwise Akaike information criterion (AIC) regression strategy.

Reweighted baseline values of second-line subjects of CELESTIAL trial are presented below in Table 27.

Table 27. Comparison of reweighted baseline characteristics of subjects enrolled in CELESTIAL (pure 2nd line) and RESORCE

Treatment (N)	CELESTIAL pure 2nd line		RESORCE
	Pure 2nd line (S1)	Pure 2nd line (S2)	As reported
	Cabozantinib (N = 187.27)	Cabozantinib (N = 303.24)	Regorafenib (N = 374)
Age under 65 (%)	54.97##	53.34##	54.97##
Female (%)	18.63	12.04	12.04
Asia geographical region (%)	37.7	22.93	37.7
ECOG status 0 (%)	65.79	65.79	65.79
Child-Pugh class A (%)	97.91	98.86	97.91
Mean duration of sorafenib treatment (months)	11.63	7.52	11.63
Extrahepatic disease (%)	71.9	71.9	71.9
Macrovascular invasion (%)	28.62	28.62	28.62
Hepatitis B aetiology (%)	37.7	37.92	37.7
Alcohol use aetiology (%)	25.31	22.78	25.31
Hepatitis C aetiology (%)	20.77	24.53	20.77
AFP > 400ng/mL (%)	43.46	43.46	43.46
White (%)	35.95	58.15	35.95

Abbreviations: AFP; alpha fetoprotein; ECOG, Eastern Cooperative Oncology Group; N, number of patients; S, scenario; ESS; Effective Sample Size.

[##]Approximate ESS values.

Statistical analysis

The baseline characteristics used for the matching procedure were selected from the preliminary set based on their potential influence on key efficacy (PFS and OS) and safety outcomes (AEs).

The NICE Decision Support Unit Technical Support Document 18 recommends justifying the choice of matching parameters by clinical expert advice and empirical identification of all prognostic variables and effect modifiers included in the weighting model. The clinical relevance of potential matching variables was justified by clinical experts on a UK advisory board meeting on the 28th June 2018 and further validated at an advisory board meeting on 31st March 2021 [4, 55]. The baseline characteristics available for matching in both trials and deemed potential effect modifiers by the clinical experts were age group, race, geographical region, ECOG performance status, Child-Pugh class, duration of prior sorafenib treatment, extrahepatic disease, macrovascular invasion, aetiology of HCC (Hepatitis B, alcohol use and Hepatitis C), and AFP level.

Additionally, effect modifiers for the primary survival endpoint, OS, are identified empirically via a stepwise AIC regression strategy. In this strategy, candidate baseline characteristics were added (or eliminated) from a regression model using a stepwise process based on the AIC. The stepwise model comparison was run in all directions (forward, backward and both) [56]. In all cases, the predictors giving the lowest AIC were gender, ECOG performance status, extrahepatic disease, macrovascular invasion and AFP level. These predictors were clinically plausible effect modifiers, except for gender as per clinical feedback received from the advisory board and hence not included for matching [55]. The baseline characteristics used for matching, and the matching scenarios considered are summarised in Table 28.

Table 28: Baseline characteristics selected for matching

Clinical expert selection (scenario 1)	Empirical analysis (scenario 2)
ECOG performance status	ECOG performance status
Baseline HCC disease per CRF (EHS and MVI)	Baseline HCC disease per CRF (EHS and MVI)
AFP level >400ng/ml	AFP level >400ng/ml
Age group	Gender
Child-Pugh class	

Clinical expert selection (scenario 1)	Empirical analysis (scenario 2)
Duration of prior sorafenib treatment	
Race	
Aetiology of HCC (Hepatitis B, alcohol use and Hepatitis C)	
Geographical region	

Abbreviations: AFP, alpha fetoprotein; CRF, case report form; ECOG, Eastern Cooperative Oncology Group; EHS, extrahepatic spread; HCC, hepatocellular carcinoma; MVI, macroscopic vascular invasion.

Safety

The estimated relative effects of cabozantinib versus placebo in the RESORCE population are found by taking weighted means of the AE outcomes in the CELESTIAL trial. These estimates have been generated using a linear model. This allows for the correct calculation of standard errors using a robust sandwich estimator [57]. The log ORs of regorafenib versus placebo are computed using the reported data on AEs. The variance of the log ORs is approximated using the delta method. The indirect comparison estimates of cabozantinib versus regorafenib are constructed in the log OR scale, using the fact that they are equal to the estimated effects (log OR) of cabozantinib versus placebo minus the estimated effects of regorafenib versus placebo in the RESORCE population.

Results

Efficacy outcomes

The selected PLD from CELESTIAL was adjusted to match aggregate data from RESORCE, survival outcomes were recalculated for each pure second-line patient in CELESTIAL using the weighted data. The pure second-line patient population from CELESTIAL had a median follow-up of 22.6 months. Table 29 presents summary statistics with 95% confidence intervals for the (weighted and unweighted) Kaplan-Meier curves fitted to the cabozantinib and regorafenib survival data. For example, for regorafenib OS at the first quartile (i.e., 75% of patients are alive), 4.9 months have elapsed. Confidence intervals for quartiles use Woodruff's method: the interval is the intersection of the horizontal line at the specified quartile with the pointwise confidence band around the survival curve [58]. This analysis suggests statistically significant differences at the 5% level for PFS but not for OS. Given the similarity between the scenarios, scenario 1 was considered the base case [58].

Table 29: Durations for endpoint Kaplan-Meier quartiles with 95% confidence intervals (in parentheses)

Treatment	PFS			OS		
	Q1 (months)	Q2 (months)	Q3 (months)	Q1 (months)	Q2 (months)	Q3 (months)
Regorafenib	1.45 (1.45-1.76)	3.19 (2.78-4.14)	6.99 (5.91-8.38)	4.90 (4.22-5.65)	10.79 (9.18-12.30)	20.96 (18.42-25.29)
Cabozantinib (unweighted pure-second line population)	2.07 (1.87-3.15)	5.52 (4.67-5.68)	9.20 (7.82-10.97)	5.91 (4.86-7.03)	11.24 (9.53-13.96)	21.85 (19.52-24.51)
Cabozantinib (weighted pure second-line population; Scenario 1)	2.37 (1.91-3.71)	5.59 (4.90-7.26)	9.56 (7.85-11.07)	5.78 (4.34-7.06)	11.37 (8.90-16.95)	22.74 (19.58-33.74)
Cabozantinib (weighted pure second-line population; Scenario 2)	2.10 (1.87-3.61)	5.55 (4.90-5.91)	9.20 (7.82-10.97)	6.21 (5.06-7.33)	11.50 (9.56-14.00)	22.05 (19.58-25.66)

Abbreviations: OS, overall survival; PFS, progression-free survival; Q, quartile.

Safety outcomes

The log OR estimates are anchored because they use the common placebo arm. An anchored log OR estimate cannot be constructed for the diarrhoea AE because it has no occurrences in the placebo arm (giving a log OR of infinity for regorafenib versus placebo). For any AEs that do not occur in a given trial arm, approximate unanchored estimates of the log ORs are performed. Palmar-plantar erythrodysesthesia is another AE for which an unanchored estimate is performed, as it does not occur in the placebo arm of CELESTIAL pure second-line.

Table 30 presents the resulting anchored AE log ORs with 95% confidence intervals, standard errors and p-values.

Table 30: log ORs, confidence intervals, std. errors and p-values for treatment-emergent grade 3/4 AEs (cabozantinib vs. regorafenib)

Adverse event	CELESTIAL data	log OR	95% CI	standard error	p-value
Increased AST	Unweighted pure 2nd line	0.89	-0.31-2.09	0.61	0.1478

Adverse event	CELESTIAL data	log OR	95% CI	standard error	p-value
	Weighted pure 2nd line (S1)	0.79	-0.47-2.06	0.65	0.2201
	Weighted pure 2nd line (S2)	0.94	-0.29-2.17	0.63	0.1352
Elevated bilirubin	Unweighted pure 2nd line	-0.55	-3.01-1.91	1.25	0.6732
	Weighted pure 2nd line (S1)	-0.25	-2.73-2.23	1.26	0.8558
	Weighted pure 2nd line (S2)	-0.21	-2.67-2.25	1.26	0.8766
Fatigue	Unweighted pure 2nd line	0.07	-1.65-1.79	0.88	0.9404
	Weighted pure 2nd line (S1)	0.09	-1.77-1.94	0.95	0.9313
	Weighted pure 2nd line (S2)	0.4	-1.35-2.14	0.89	0.671
Hypertension	Unweighted pure 2nd line	1.73	-0.45-3.91	1.11	0.1207
	Weighted pure 2nd line (S1)	2.1	-0.1-4.3	1.12	0.0611
	Weighted pure 2nd line (S2)	1.72	-0.47-3.9	1.11	0.1239
Diarrhoea (unanchored)	Unweighted pure 2nd line	1.55	0.8-2.3	0.38	0.0001
	Weighted pure 2nd line (S1)	1.74	1-2.48	0.38	<0.0001

Adverse event	CELESTIAL data	log OR	95% CI	standard error	p-value
	Weighted pure 2nd line (S2)	1.68	0.94-2.43	0.38	<0.0001
Palmar-plantar erythrodysesthesia (unanchored)	Unweighted pure 2nd line	0.3	-0.17-0.77	0.24	0.2103
	Weighted pure 2nd line (S1)	0.05	-0.4-0.5	0.23	0.848
	Weighted pure 2nd line (S2)	0.3	-0.15-0.76	0.23	0.1934

Abbreviations: AE; adverse event; CI, confidence interval; OR, odd ratio; S, scenario.

Sensitivity analysis of the anchored MAIC

In order to assess differences between cabozantinib and regorafenib OS and PFS, the proportional hazards assumption was assessed.

Figure 13 presents the log-cumulative hazard plot of weighted cabozantinib (Scenario 1) versus regorafenib for the PFS outcome. The curves remain parallel till after month 10 where the curves eventually cross. This would suggest that the proportional hazards assumption is not satisfied for the PFS outcome however there are low patient numbers generating the tail of these curves. The plot of the scaled Schoenfeld residuals (Figure 14) shows a degree of flatness however the Grambsch-Therneau test has a p-value of 0.0002 which indicates a non-zero slope.

Figure 13: PFS log-cumulative hazard plot for weighted cabozantinib (Scenario 1) versus regorafenib

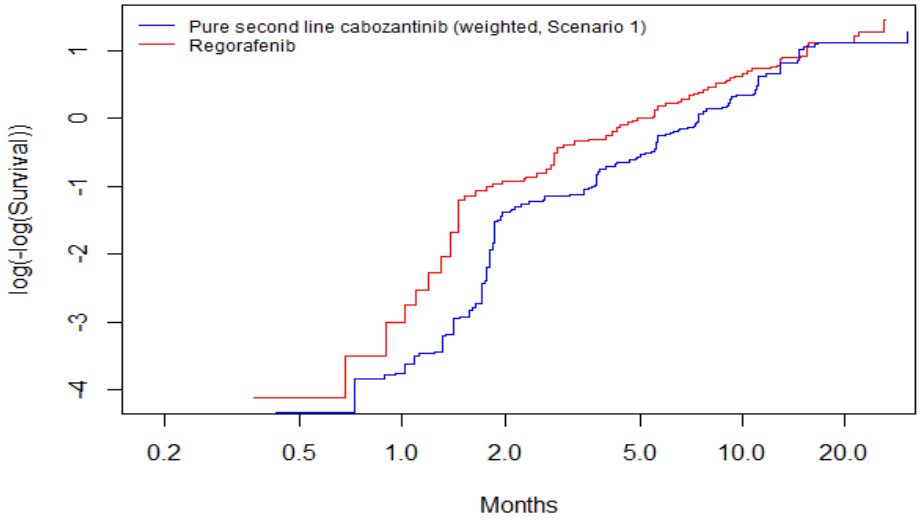


Figure 14: Scaled Schoenfeld residuals for PFS for weighted (Scenario 1) pure second-line cabozantinib versus regorafenib

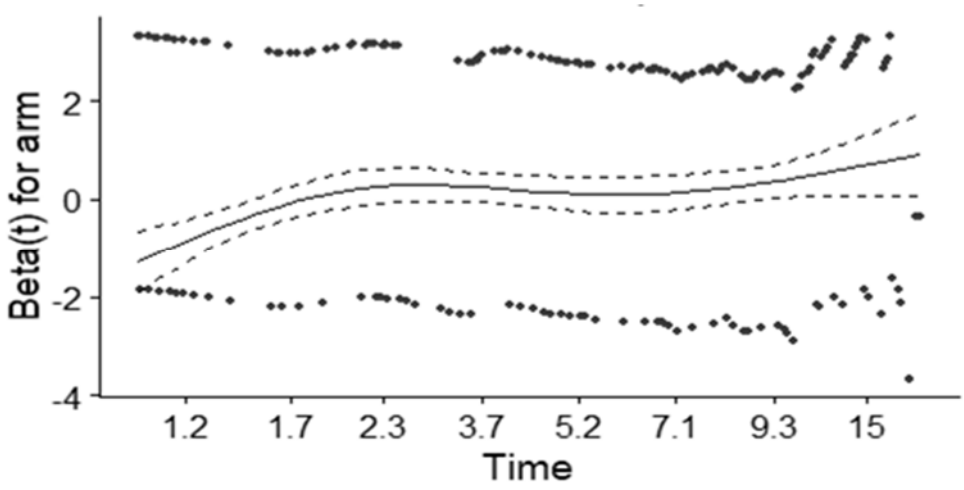


Figure 15 presents the corresponding log-cumulative hazard plot for the OS outcome. The OS curves cross at several instances. These intertwined curves suggest that the OS outcomes of the groups are similar. Similar to PFS, the plot of the scaled Schoenfeld residuals (Figure 16) shows a degree of flatness however the Grambsch-Therneau test (p-value 0.0029) indicates a non-zero slope as well.

Figure 15: OS log-cumulative hazard plot for weighted cabozantinib (Scenario 1) versus regorafenib

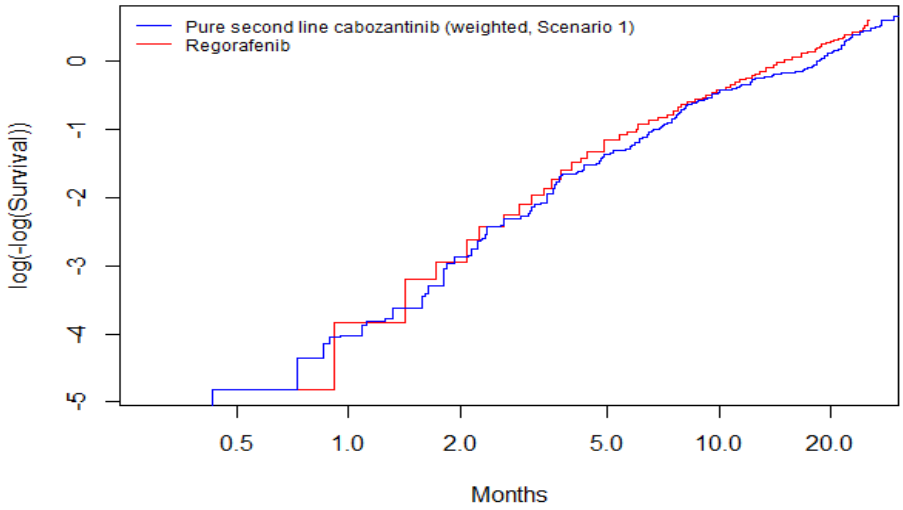
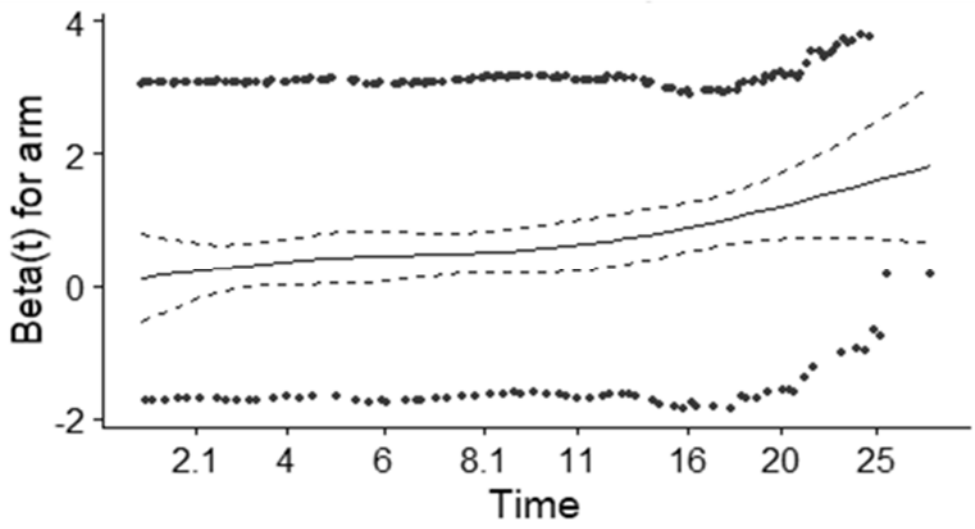


Figure 16: Scaled Schoenfeld residuals for PFS for weighted (Scenario 1) pure second-line cabozantinib versus regorafenib



Given the uncertainty of the proportional hazard assumption for both endpoints, a range of models were explored which would further assess the uncertainty of whether there was any difference in treatment effect between cabozantinib and regorafenib, as summarised below:

- An anchored analysis assuming that the proportional hazards assumption holds between cabozantinib and regorafenib. This analysis uses a constant Cox HR of weighted CELESTIAL data and RESORCE to generate a hazard ratio between cabozantinib and regorafenib;
- An anchored analysis assuming that the proportional hazards assumption does not hold. This analysis explores if there is any difference in treatment effect

emerging between cabozantinib and regorafenib over time. This is conducted by generating time-varying hazard ratios from hazard profiles of fitted parametric models to the weighted CELESTIAL and RESORCE data;

- An unanchored analysis comparing the treatment effect by using fitted parametric models to weighted cabozantinib and regorafenib data.

Anchored analysis using constant HR

The results of an anchored comparison between cabozantinib and regorafenib using a constant hazard ratio are shown in Table 31. The hazard ratio of cabozantinib versus regorafenib shows a point estimate that favours PFS for cabozantinib, while the opposite for OS. Both of these results are not statistically significant.

Table 31. Results of anchored comparison using a constant hazard ratio

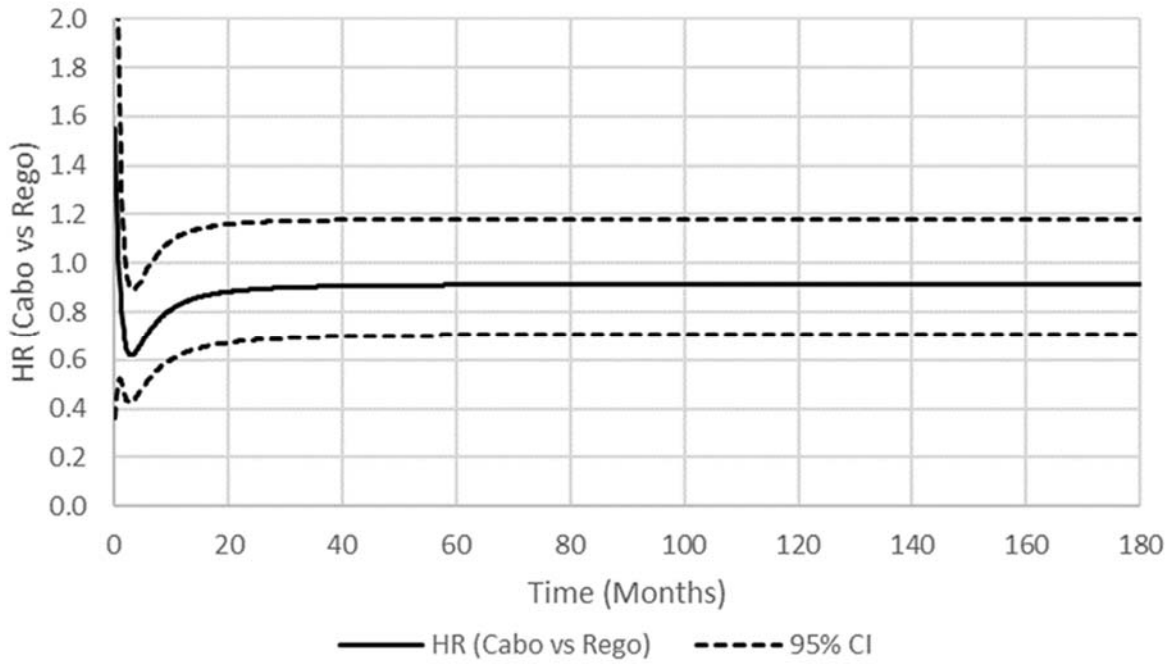
Endpoint: relative effect measure	Weighted CELESTIAL: Cabozantinib versus placebo HR (95% CI)	RESORCE: Regorafenib versus placebo HR (95% CI)	Cabozantinib versus regorafenib HR (95% CI)
Overall survival	0.73 (0.54, 0.99)	0.63 (0.50, 0.79)	1.15 (0.79, 1.69)
Progression-free survival	0.36 (0.28, 0.48)	0.46 (0.37, 0.56)	0.79 (0.56, 1.11)

Abbreviations: CI, confidence interval; HR, hazard ratio.

Anchored analysis using time-varying HR

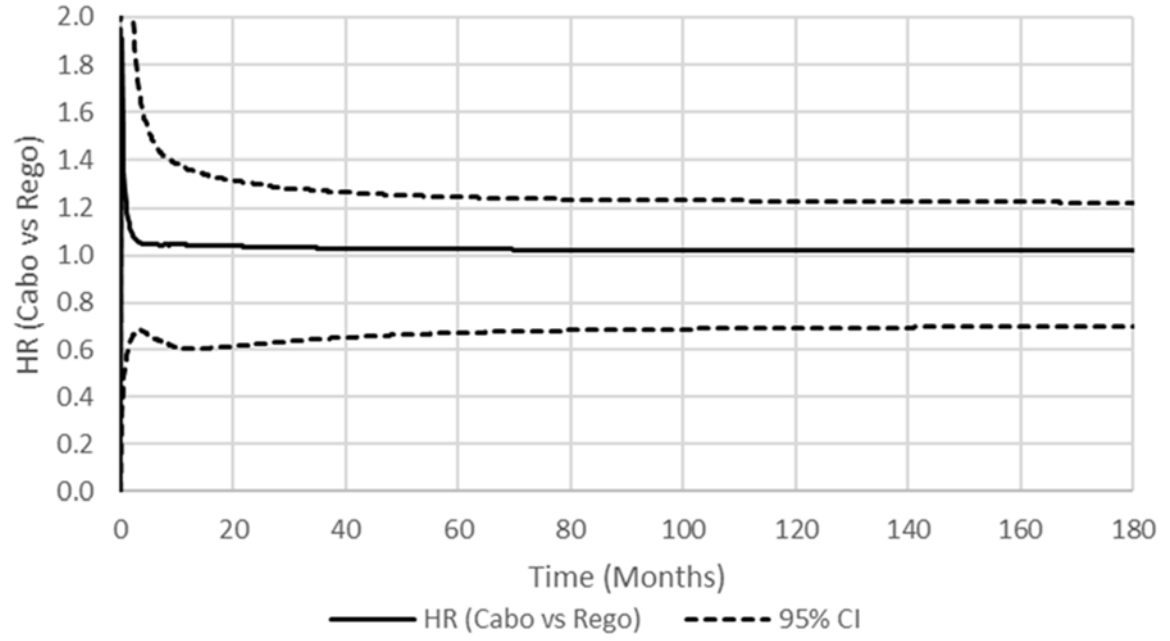
The result of the anchored analysis using time-varying hazard ratios generated from the log-logistic model is shown in Figure 17 for PFS and in Figure 18 for OS. For both endpoints the log-logistic model was the best fitting by AIC and Bayesian information criterion (BIC); however, the other standard parametric models were tested and the results are shown in Appendix I. The results across the models show that over time, the hazard ratio is not statistically different from 1, indicating no difference in treatment effect. Furthermore, the hazard ratio is generally seen to be constant and near 1 as the treatment effect is extrapolated which suggests equivalence in treatment effect over time. Similar to the constant hazard ratio analysis, the point estimate shows conflicting direction of treatment benefit as there is a benefit for cabozantinib for PFS but a benefit for regorafenib for OS.

Figure 17: Log-logistic model for time-varying hazard ratio of cabozantinib versus regorafenib for progression-free survival endpoint



Abbreviations: Cabo, cabozantinib; CI, Confidence interval; HR, Hazard ratio; Rego, regorafenib.

Figure 18: Log-logistic model for time-varying hazard ratio of cabozantinib versus regorafenib for overall survival endpoint

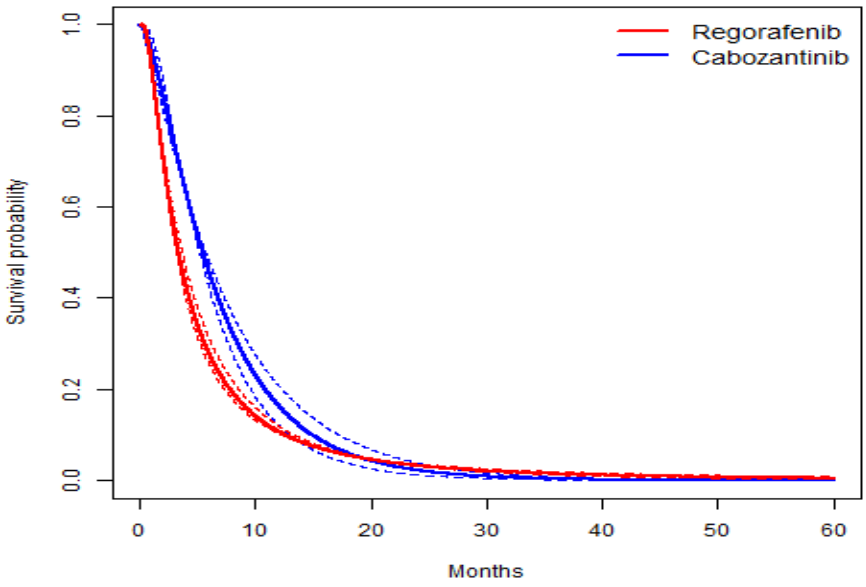


Abbreviations: Cabo, cabozantinib; CI, confidence interval; HR, hazard ratio; Rego, regorafenib.

Unanchored analysis using independent parametric models

The results of the unanchored analysis for PFS are shown in Figure 19 using a generalised gamma model. The generalised gamma was the best fitting model by AIC and BIC. Confidence intervals were produced by simulating a large bootstrap-like sample from the asymptotic normal distribution of the maximum likelihood estimates of the parameters [59]. In total, 100,000 random samples were drawn to ensure that the recovered mean and median survival times were stable to two decimal places through different runs. The models show a statistically significant benefit for cabozantinib until approximately 1 year when the PFS curves show little difference for the rest of the time horizon. Cabozantinib has a larger point estimate for mean PFS than regorafenib (7.17 vs. 6.04 months) and higher median PFS (5.49 vs. 3.39).

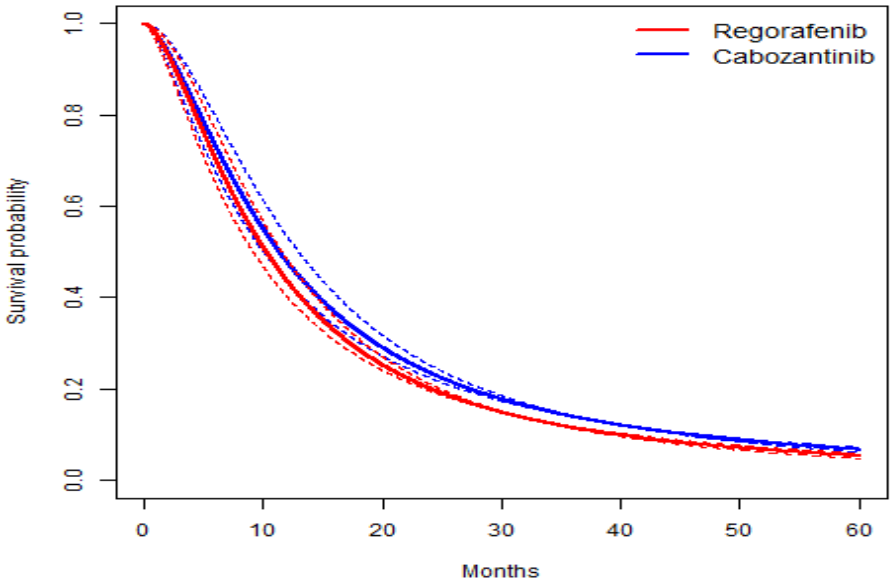
Figure 19: Unanchored results for PFS



Abbreviations: PFS, progression-free survival.

The results of the unanchored analysis for OS are shown in Figure 20 using a generalised gamma model. The OS curves show a large amount of overlap until year 1 when cabozantinib begins to show a relatively small benefit over regorafenib. Cabozantinib has a larger point estimate for mean OS (24.65 vs. 21.17 months) and a higher median OS (11.40 versus 10.29 months).

Figure 20: Unanchored results for OS



Abbreviations: OS, overall survival.

B.3.10.4 Discussion and conclusions of indirect treatment comparisons

There was presence of between-study heterogeneity among CELESTIAL and RESORCE trials, namely the increased tolerability of patients to sorafenib in RESORCE and the inclusion of third-line patients in the CELESTIAL population. Despite the different populations, the Bucher approach showed that for the point estimates, OS favoured regorafenib but PFS slightly favoured cabozantinib. None of these results are statistically significant. The proportional hazards assumption did not hold for OS; therefore, the treatment effect may not be representative as a constant hazard ratio.

When adjusting for population differences through the MAIC, the anchored analysis showed that cabozantinib has a higher point estimate than regorafenib for PFS; however, regorafenib was associated with higher OS (point estimate) than cabozantinib. None of these results were statistically significant. Relaxing the proportional hazards assumption through the time-varying hazard ratio analysis showed no significant difference for the treatment effect over time. The unanchored MAIC as a scenario analysis to the anchored approach showed that cabozantinib may achieve a similar OS and prolonged PFS compared with regorafenib. The improvement in PFS was statistically significant in favour of cabozantinib

A previously published MAIC study using real world evidence (RWE) for regorafenib showed similar results to that provided in this submission. The Casadei Gardini et al. analysis used data from 278 patients who received regorafenib as a second-line therapy after previous treatment with sorafenib for unresectable HCC. This group of patients also included those intolerant to sorafenib as well as tolerant, whereas the RESORCE trial only had sorafenib tolerant patients. Published aggregate data for the subgroup of CELESTIAL patients who received sorafenib as the only prior therapy were used in the analysis for cabozantinib data [60]. This methodology estimates the effect of the regorafenib treatment in the patient population that received cabozantinib. The results found cabozantinib to have a statistically significant benefit over regorafenib in terms of PFS in all prior sorafenib patient populations [HR 0.50 (0.41-0.62)]. It also found a benefit in terms of OS with point estimates in favour of cabozantinib versus regorafenib [HR 0.83 (0.62-1.09)] but this was not statistically significant [61]. Other network meta-analyses (NMAs) that have been conducted and reported in the literature have similarly found no statistically significant difference between the two treatment options in terms of survival or safety endpoints. The OS and PFS results are summarised in Table 32

Table 32. Results from ITCs conducted in the literature

Study	Overall survival (HR 95% CI)	Progression-free survival (HR 95% CI)
Wang et al.2020 [62]	Rego vs Cabo: 0.82 (0.63-1.1)	Rego vs Cabo: 1.1 (0.80-1.4)
Bakouny et al. 2018 [63]	Rego vs Evero: 0.60 (0.44-0.51) Cabo vs Evero: 0.72 (0.55-0.95)	Rego vs Evero: 0.46 (0.35-0.62) Cabo vs Evero: 0.47 (0.36-0.63)
Sonbol et al. 2020 [64]	Rego vs Cabo: 0.82 (0.62-1.07)	Rego vs Cabo: 1.04 (0.79-1.36)
Park et al. 2021 [65]	Cabo vs Rego: 0.96 (0.54-1.68)	-
Casadei Gardini et al. 2021 [61]	Cabo vs Rego: 0.83 (0.62-1.09) Subgroups: Prior sorafenib < 3 months: Cabo vs Rego: 0.68 (0.39-1.16) Prior sorafenib 3 to 6 months: Cabo vs Rego: 0.66 (0.42-1.02)	Cabo vs Rego: 0.50 (0.41-0.62) Subgroups: Prior sorafenib < 3 months: Cabo vs Rego: 0.33 (0.21-0.50) Prior sorafenib 3 to 6 months: Cabo vs Rego: 0.53 (0.37-0.75)

Study	Overall survival (HR 95% CI)	Progression-free survival (HR 95% CI)
	Prior sorafenib > 6 months: Rego vs Cabo: 0.89 (0.52-1.51)	Prior sorafenib > 6 months: Cabo vs Rego: 0.60 (0.38-0.94)

Abbreviations: Cabo: cabozantinib; Evero, everolimus; HR, hazard ratio; ITC, indirect treatment comparison; Rego: regorafenib.

The AE analysis using a Bucher approach, showed different point estimates for AEs that were able to be analysed through the Bucher approach, but the results were not significant. When using the MAIC methodology, only diarrhoea shows statistically significant differences at the 5% level. However, this estimate is unreliable because the grade 3/4 treatment-emergent AE only occurs twice for the CELESTIAL placebo arm and never occurs for the RESORCE placebo arm. The patients in RESORCE were tolerant to sorafenib and this would reduce the occurrence of grade 3/4 treatment-emergent diarrhoea. Some of the anchored log ORs are very large (e.g., the estimates for hypertension are close to 2), probably a result arising from very small counts in the data, particularly in the CELESTIAL placebo arm, which make the estimates unprecise and drive them upward.

The RWE data shows that cabozantinib has a similar toxicity profile to that observed in the CELESTIAL trial with certain grade 3+ AEs of interest occurring closer to that of the numbers reported in the RESORCE trial [66, 67].

In conclusion, the ITC results suggest that cabozantinib has comparable or greater clinical efficacy and similar tolerability compared to regorafenib, thus justifying the approach of a cost-comparison analysis for cabozantinib versus regorafenib as it is intended for interventions that demonstrate similar or greater health benefits than technologies already recommended by NICE in technology appraisal guidance.

B.3.10.5 Uncertainties in the indirect and mixed treatment comparisons

The population differences between the trials introduced bias into the Bucher analysis. Therefore, a MAIC was conducted to reduce the impact of these variables on the results. The effective sample size for the MAIC remained large with 265.53 for scenario 1 and 452.31 for scenario 2. There were some large, rescaled weights in scenario 1 with a maximum of 9.21 but scenario 1 matched with more characteristics that are considered to be important effect modifiers by the clinical experts, and which

differ considerably across trials (e.g. duration of prior sorafenib treatment and geographical region). The two scenarios produced similar results.

A negative outcome control was conducted as a form of validation. This compared the weighted placebo arm of CELESTIAL and the placebo arm of RESORCE. The MAIC can balance observed patient characteristics but there is still the potential for residual confounding due to unobserved differences between trials. The recovered HR for OS (CELESTIAL placebo vs. RESORCE placebo) was 0.87 (95% confidence interval 0.67-1.15; p-value 0.326). For Scenario 2, the estimated HR for OS is 0.88 (95% confidence interval 0.68-1.14; p-value 0.326). In both cases, the HR was close to one. The recovered HR for PFS was 0.69 (95% confidence interval 0.55-0.87; p-value 0.00158). For Scenario 2, the estimated HR for PFS was 0.72 (95% confidence interval 0.58-0.90; p-value 0.00328). This would suggest that, even after matching, there remains important cross-trial differences in the placebo arms. There is therefore some sort of residual imbalance impacting the PFS outcomes. This adds uncertainty to any superiority claim in terms of PFS benefit for cabozantinib over regorafenib and thus equal equivalence is assumed in this submission as a conservative assumption.

The uncertainty regarding the proportion hazards assumption was explored by investigating the trend of the hazard ratio over time between cabozantinib and regorafenib. The time-varying hazard ratio analysis was able to show that there was no significant difference for the treatment effect over time. A further sensitivity was conducted by not using the hazard ratio to represent the treatment effect but instead fit independent curves to the cabozantinib and regorafenib arms. This showed similar or better treatment effect for cabozantinib which is in line with the conservative assumption of equal equivalence between treatments.

B.3.11 Conclusions about comparable health benefits and safety

Cabozantinib is indicated as monotherapy for the treatment of HCC in adults who have previously been treated with sorafenib. The proposed positioning of cabozantinib as a treatment option after prior treatment with sorafenib offers an alternative treatment option to a UK patient population with poor prognosis where there is only one other treatment option currently recommended by NICE. For these patients, cabozantinib offers an additional treatment option, including patients intolerant to sorafenib.

Cabozantinib is an oral multi-targeted inhibitor of RTKs that delivers significantly extended survival and delayed disease progression in patients with advanced HCC who have received prior therapy. This is supported by a robust, high quality phase 3 clinical programme as well as with indirect evidence versus regorafenib (the comparator in this submission) in the form of a Bucher ITC and MAIC.

The CELESTIAL trial was an international, randomised, double-blind, placebo-controlled, phase III trial. In the CELESTIAL trial, at the cut-off date for the second interim analysis of OS (01 June 2017), there was high maturity with a total of 484 deaths (78% actual information fraction) reported. The trial shows cabozantinib significantly reduced the risk of death by 24% compared with placebo and significantly reduced the risk of disease progression/death by 56% compared with placebo. Cabozantinib was associated with a significantly higher ORR than placebo. Consequently, cabozantinib was also associated with a lower rate of PD compared with placebo (21% versus 55%).

The benefits of cabozantinib were accompanied by a manageable safety profile, as illustrated by patients in the cabozantinib group staying on treatment for almost twice as long as those in the placebo group (3.8 versus 2.0 months). Many AEs were as expected in patients with pre-treated advanced HCC, reflected by their high frequency in both the placebo and cabozantinib groups. The most frequently reported AEs in the cabozantinib group were typical of those with VEGFR-TKI therapies such as regorafenib [41] and consistent with the known safety profile of cabozantinib in patients with advanced renal cell carcinoma [11]. This is further supported by RWE studies such as those discussed in Table 32.

Conclusions from the evidence of the cabozantinib phase 3 clinical trial programme are supplemented by indirect comparisons designed to compare cabozantinib to regorafenib which was not included in the trial programme, but is relevant to National Health Service (NHS) clinical practice. Across these analyses, cabozantinib demonstrated comparable efficacy and a similar safety profile to regorafenib. This was shown through the conflicting direction of treatment benefit of the point estimates for OS and PFS. The confidence intervals showed that this was not statistically significant for OS. However, for PFS certain analyses showed a statistically significant treatment benefit for cabozantinib over regorafenib. Time-varying hazard ratio analyses showed

that there was no divergence in treatment effect between the treatments over time. Additionally, evidence from the ITCs confirmed the rates of AEs are comparable across treatments.

There are existing uncertainties in the ITC which have been explored through a range of modelling techniques designed to establish the comparative treatment effect between cabozantinib and regorafenib. There was evidence to suggest that all the heterogeneity between the trials could not be accounted for, thus a conservative assumption of equal efficacy is assumed, especially for the PFS endpoint as this favours cabozantinib. This assumption is in line with clinical expert feedback received during an advisory board [4] and responses received by NICE from professional bodies to the scoping consultation [6].

B.3.12 Ongoing studies

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

B.4 Cost-comparison analysis

B.4.1 Changes in service provision and management

Cabozantinib is an oral tablet that is administered once daily and it is not anticipated to require any changes to current service provision and management of the target patient population eligible for treatment in the NHS England setting. This was also indicated in the responses received by NICE from professional bodies to the scoping consultation [6].

No differences in resource use are anticipated between cabozantinib and regorafenib (Section B.4.2.3 and B.4.2.4). A cost and resource SLR was conducted but did not identify any studies that would indicate differential health care resource use between the treatments.

B.4.2 Cost-comparison analysis inputs and assumptions

B.4.2.1 Features of the cost-comparison analysis

A cost-comparison analysis was conducted to evaluate the cost to the NHS of using cabozantinib versus regorafenib for treating adults with advanced HCC who have had sorafenib. A simple economic model was developed in Microsoft Excel to facilitate the comparison.

As introduced in Section B.1, regorafenib was selected as the appropriate comparator because:

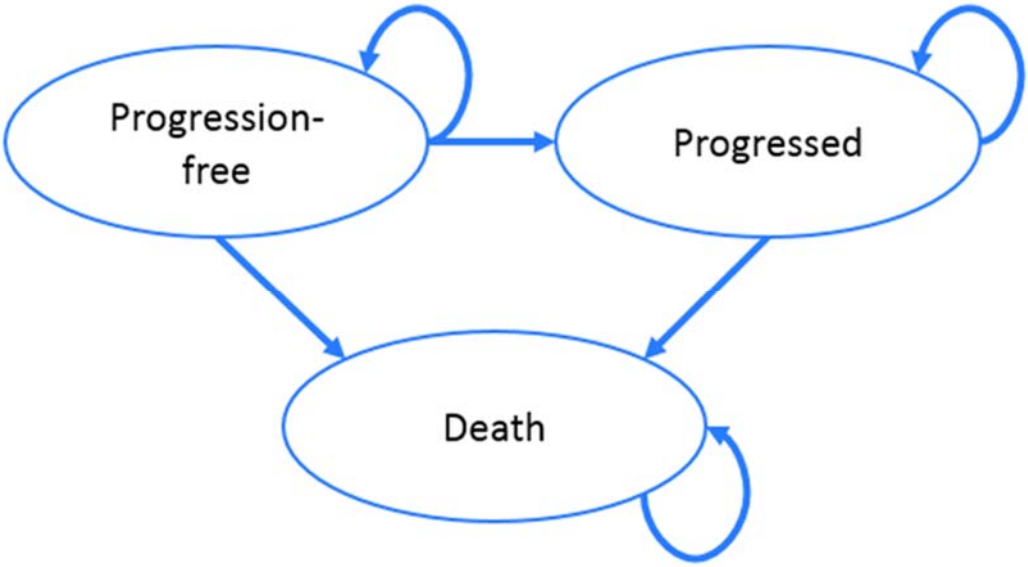
- It is recommended by NICE for its licensed indication, adults with advanced unresectable HCC who have previously been treated with sorafenib. Cabozantinib has the same licensed indication and Ipsen are seeking the same positioning as the NICE recommendation for regorafenib. Since regorafenib is the only approved subsequent therapy for use after sorafenib, it is assumed to have a majority market share in this indication. This is supported by clinical experts estimation of regorafenib market share within the indication [4]
- In post sorafenib patients eligible for treatment in the second and third-line setting, regorafenib is used in clinical practice. This is following the approval of atezolizumab plus bevacizumab in first-line, where sorafenib is now positioned

as a second-line treatment option in addition to its use in first-line [5]. Similarly, to second-line use, patients eligible for regorafenib in third-line are restricted to patients with ECOG performance status of 0 or 1 for which BSC is not a relevant treatment option. Therefore, regorafenib is the only comparator in this setting as patients would not be fit enough to receive chemotherapy. Consequently, regorafenib is the only appropriate comparator used in clinical practice which should form the basis for decision making.

In line with ERG and committee feedback on TA555 for regorafenib, a 15-year time horizon was adopted in the analysis to capture costs over a sufficient length of time. Only drug acquisition costs were considered in the base case analysis as all other costs were assumed equal given the equal efficacy and method of administration of the treatments. The equal efficacy assumption was relaxed in scenario analyses where the cost of the drug-specific toxicity profiles was taken into account. Costs were not discounted in line with the user guide for cost-comparison for FTA [68].

The model calculates the incremental cost by calculating the product of the mean time on treatment and the drug pack price for each treatment. Given the equal efficacy assumption, both cabozantinib and regorafenib were assumed to have the same time on treatment. The mean time on treatment in the model was estimated by calculating a 15-year restricted mean of the PFS curve since patients are treated to progression. In line with previous models in advanced HCC (Table 5), patients follow a 3-health state model that has progression-free, progressed and death health states, as illustrated in Figure 21. Patients remain on treatment in the progression-free health state hence the PFS curve is an appropriate estimator of time on treatment and is the only efficacy outcome required to inform the cost-comparison.

Figure 21: 3 health state model structure diagram



The PFS curve was generated by fitting independent parametric models to the IPD from the CELESTIAL trial and extrapolating to the end of the time horizon as per NICE TSD 14 [69]. The PFS curve was bounded by a parametric model fitted to OS, though no curve crossing was observed. Figure 22 shows the parametric models fitted to the PFS IPD for cabozantinib from the CELESTIAL ITT. The statistical fit is shown in Table 33. All models had a good visual fit and the generalised gamma and log-logistic models were the best fitting by AIC and BIC. The top two models were within approximately 3 AIC and BIC suggesting a similar statistical fit. The long-term progression-free survival extrapolations from the parametric models were presented to three clinical experts. Based on their clinical experience, these experts estimated that PFS at 2 years and 4 years will be 5% and 1%, respectively [55]. The log-logistic curve was used in the base case as the 4-year PFS probability is 1% for the log-logistic model compared to 0% for the generalised gamma. This resulted in a 15-year restricted mean of [redacted]

Figure 22: PFS cabozantinib parametric fits

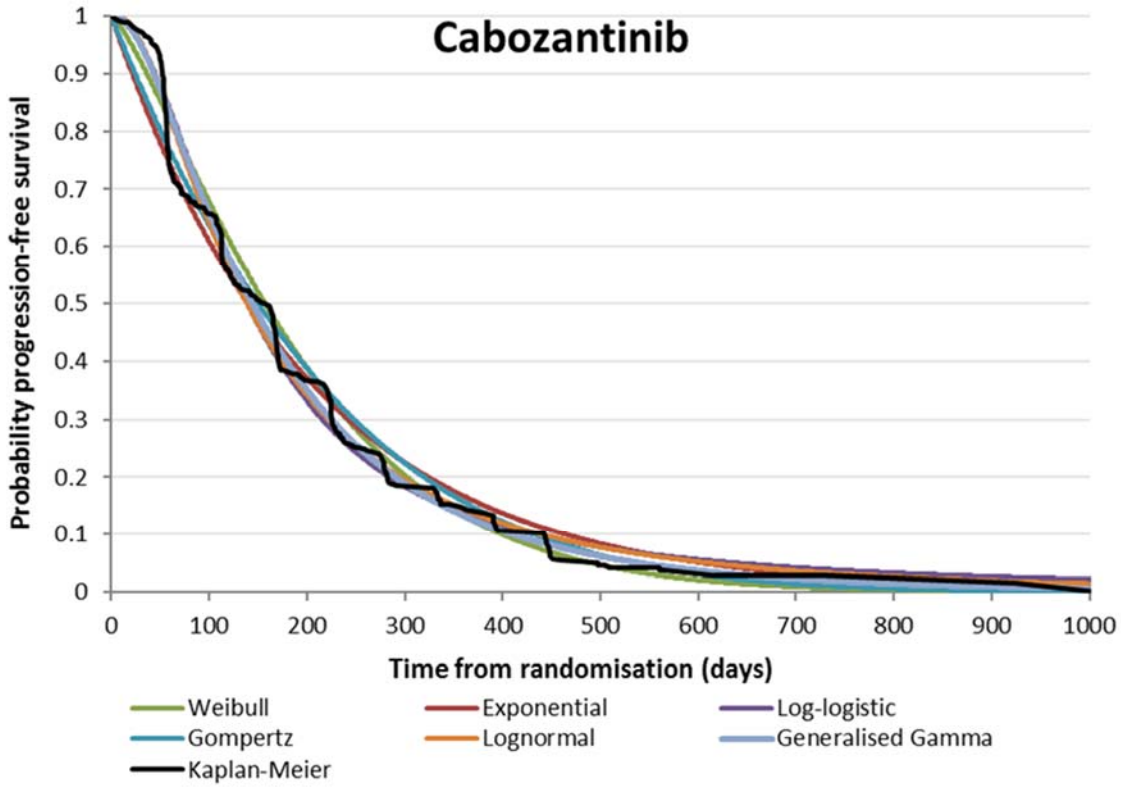


Table 33: AIC and BIC statistics for PFS parametric fits

Model	AIC	BIC
Exponential	1079.16	1083.32
Weibull	1046.11	1054.41
Log-logistic	1025.94	1034.25
Gompertz	1073.84	1082.14
Lognormal	1027.69	1036.00
Generalised gamma	1022.70	1035.16

Abbreviations: AIC, akaike information criterion; BIC, bayesian information criterion.

Patients treated with cabozantinib are assumed to receive 60mg every day, whereas patients treated with regorafenib receive 160mg every day for 3 weeks in a 4-week cycle as per the licensed recommended dose. The cost of drug wastage is included for both treatments such that [redacted] packs of cabozantinib tablets are costed for [redacted] cabozantinib cycles ([redacted] cycles is equal to treatment duration, [redacted] months divided by cabozantinib cycle length, 30 days, [for conversion of months to days in the model, 1 month is assumed to be 365.25/12 days]). The number of treatment cycles for regorafenib is [redacted] cycles which requires [redacted] packs of regorafenib tablets ([redacted] cycles is equal to treatment duration, [redacted] months divided by regorafenib cycle length, 28 days).

It should be noted a regorafenib cycle between 0.75 and 1.0 costs the same as 1.0 treatment cycle since the last week of a treatment cycle accrues £0 cost.

B.4.2.2 Intervention and comparator's acquisition costs

Table 34 presents a summary of the key inputs, assumptions and acquisition costs included for cabozantinib and regorafenib. As cabozantinib has a treatment cycle length of 30 days compared to 28 days for regorafenib, the number of packs required for treatment is less with cabozantinib for a sufficiently long treatment duration, e.g., 12.18 packs are required for cabozantinib for a year of treatment compared to 13.04 packs for regorafenib when no drug wastage is assumed.

Table 34: Key inputs to quantify the acquisition costs of cabozantinib and regorafenib

	Cabozantinib	Regorafenib
Pharmaceutical formulation	60mg oral tablet	40mg oral tablet
(Anticipated) care setting	Hospital prescription/supply	Hospital prescription/supply
Acquisition cost (excluding VAT)	List price of £5,143.00 per pack of 30 x 60mg tablets Ipsen proposed a confidential PAS which results in the price of [redacted] per pack Average cost per course of treatment over a 15 year time horizon: [redacted] calculated as list price (£5,143) x number of treatment cycles ([redacted]). Number of treatment cycles is equal to rounded up value to the nearest cycle, of average time on treatment in order to account for drug wastage ([redacted] months) / treatment cycle length (30 days)	List price of £3,744.00 per pack of 84 x 40mg tablets [14]
Method of administration	Oral	Oral
Doses	60mg dose per administration	160mg dose per administration
Dosing frequency	Cabozantinib is administered once per day	Regorafenib is administered once per day for the first 3 weeks of a 4 week cycle
Dose adjustments	N/A	N/A
Average length of a course of treatment	Average time on treatment: [redacted] months over a 15-year time horizon This is the modelled average time on treatment from the extrapolated PFS curve using restricted mean of 15 years [70]	
Average cost of a course of treatment over a 15-year time horizon (acquisition costs only) including drug wastage	[redacted]. calculated as PAS price ([redacted]) x number of treatment cycles ([redacted]). Number of treatment cycles is equal to rounded up value to the nearest cycle, of average time on	[redacted] calculated as list price (£3,744) x number of treatment cycles taking into account drug holiday ([redacted]). Number of treatment cycles is equal to rounded up value to the nearest cycle, of average time on

	Cabozantinib	Regorafenib
	treatment in order to account for drug wastage ([redacted] months) / treatment cycle length (30 days)	treatment in order to account for drug wastage ([redacted] months) / treatment cycle length (28 days) The last week of the regorafenib treatment cycle is calculated as £0 cost as no regorafenib doses are administered
Annual drug acquisition costs of treatment for a 1 year treatment duration including drug wastage	£66,859 with list price, 13 packs costed as 1 year treatment duration requires 12.18 packs of 30 x 60mg tablets [redacted], 13 packs costed as 1 year treatment duration requires 12.18 packs of 30 x 60mg tablets	£52,416, 14packs costed as 1 year treatment duration requires 13.04 packs of 84 x 40mg tablets
(Anticipated) average interval between courses of treatment	N/A – continuous treatment	
(Anticipated) number of repeat courses of treatment	N/A	N/A

Abbreviations: N/A, not applicable; VAT: value added tax

B.4.2.3 Administration and monitoring costs

As previously outlined, cabozantinib and regorafenib are administered orally which incurs £0 cost. Therefore, no administration costs were included in the analysis.

Cabozantinib requires no additional monitoring above that carried out currently for HCC. On this basis, no differences in resource use between cabozantinib and regorafenib are expected and hence such cost components are excluded from the analysis.

B.4.2.4 Adverse reaction unit costs and resource use

As reported in Section B.3.10, results of the ITC analyses for AEs indicated that the incidence of AEs associated with the use of cabozantinib and regorafenib are similar. Therefore, it is assumed that the costs associated with treating AEs would be similar for both therapies, and any difference would be negligible, and thus, AE costs were not included in the base case. A scenario tested the effect of including a different toxicity profile for cabozantinib and regorafenib. This was calculated as a one-off cost using the incidence of an AE multiplied by the respective cost. The grade 3+ treatment-related AE incidences from the MAIC used in the model are shown in Table 35. These estimates for the incidence with cabozantinib had high uncertainty due to a low number of events available for analysis as discussed in Section B.3.10. This results in some AEs, such as hypertension, having a large point estimate.

Table 35. AE grade 3 or more incidences included in scenario analysis

Adverse Events	Cabozantinib Incidence %	Regorafenib Incidence %
Palmar-plantar erythrodysaesthesia syndrome	13.2	12.6
Hypertension	55.2	13.1
Elevated aspartate aminotransferase	10.6	5.1
Fatigue	7.0	6.4
Diarrhoea	12.3	2.4
Elevated bilirubin	5.3	6.7

Abbreviations: AEs, adverse events;
Source: Bruix et al,2017 [41].

The costs of AEs have been drawn from previous NICE appraisals such as the regorafenib appraisal (TA555); however recent discussions with two clinical experts [71] have demonstrated that they are now familiar with the AE profiles of TKIs such

that most grade 3 AEs included in the scenario analysis can be managed via temporary cessation of treatment, dose reduction and supportive therapies. These AEs can often be managed via telephone discussion without the need for the patient to be seen in a hospital setting. The only grade 3 AE that clinical experts thought would warrant hospital admission would be grade 3 diarrhoea. Thus the costs of managing AEs are in in reality likely to be lower. The grade 3+ treatment-related AE costs used in the model are shown in Table 36.

Table 36: Adverse event costs

Adverse event	Cost per episode	Code, Details	Reference
Diarrhoea	£629.69	FD10K Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 6-10 - non-elective short stay	NHS reference costs 2019/20 [72]
Aspartate aminotransferase increase	£0.00	-	Based on the assumptions: regular blood tests (already considered under health state management costs)
Hypertension	£638.81	EB04Z Hypertension – Total HRG	NHS reference costs 2019/20 [72]
Fatigue	£63.45	Based on cost included in sorafenib NICE submission [39]	Inflated using PSSRU 2021 [73]
Palmar-plantar erythrodysaesthesia syndrome	£420.66	JD07J Skin Disorders without Interventions, with CC score 2-5 - non-elective short stay	NHS reference costs 2019/20 [72]
Elevated bilirubin	£0.00	-	Based on the assumptions: regular blood tests (already considered under health state management costs)

Abbreviations: NHS, National Health Service; PSSRU, Personal Socaial Services Research Unit.

B.4.2.5 Clinical expert validation

All of the parameters and assumptions, including the equivalence assumption, that were applied in the cost-comparison model were validated by a clinical expert advisory board [4]. Once the model was finalised, it was validated by internal modellers. A programmer (other than the one that built the model) reviewed all formulae and labelling in the model.

B.4.2.6 Uncertainties in the inputs and assumptions

A summary of the inputs used in the cost-comparison analysis are summarised in Table 37 and all of the key assumptions are presented in Table 38.

Table 37: Summary of model inputs

Input	Cabozantinib	Reference
Time horizon (years)	15	NICE FTA user guide [68]
Discount rate	0	NICE FTA user guide [68]
Time on treatment (mean)	[redacted]	Parametric survival analyses of CELESTIAL ITT population [70]
Costs (Cabozantinib)		
Cost per pack (List price) **Cost per pack (PAS price)**	£5,143.00 [redacted]	NICE BNF 2022 [14] Ipsen
Costs (Regorafenib)		
Cost per pack (List price)	£3,744.00	NICE BNF 2022 [14]

Abbreviations: BNF, British National Formulary; FTA, fast track appraisal; ITT, Intention to treat; NICE, National Institute of Health and Care Excellence; PAS, patient access scheme.

Table 38: Key assumptions of the analysis

Assumption	Rationale for assumption
Patients are assumed to remain on treatment till progression which is the same for cabozantinib and regorafenib	Cabozantinib and regorafenib are assumed equal efficacy therefore the time on treatment is the same
The only difference in costs are due to drug acquisition costs	Administration costs are £0 as the treatments are administered orally. Adverse events and monitoring costs are equivalent between cabozantinib and regorafenib due to equal efficacy
Drug wastage costs are included	This is a conservative assumption that there will be no efficiencies in minimising drug wastage in clinical practice. This assumption was used in TA555 [2]

B.4.3 Base case results

In the analysis presented below, the cabozantinib PAS price is compared to the regorafenib list price. Given the confidentiality of PAS prices, a cost-comparison analysis based on the cabozantinib PAS price and the regorafenib PAS price was not feasible.

Table 39 presents the base case results for a 15-year time horizon. Results show that cabozantinib can be considered a cost-saving option compared to regorafenib for the

treatment of adults with advanced HCC who have previously been treated with sorafenib. The drug acquisition costs per person over the 15-year time horizon was estimated to be [redacted] and [redacted] for cabozantinib PAS and regorafenib list price, respectively. This equates to a total cost savings of [redacted] per patient over a 15-year period.

Table 39: Base case results: 15-year time horizon

Technologies	Total costs*
Cabozantinib (PAS price)	[redacted]
Regorafenib (List price)	[redacted]
Difference	[redacted]

Abbreviations: PAS, patient access scheme.
* Drug acquistion costs were the only component considered for reasons described in Section B.4.2

B.4.4 Sensitivity and scenario analyses

Scenario analyses were performed to evaluate the sensitivity of the model results. The following scenarios were conducted:

- Varying the average time on treatment by 20%
- Relaxing the equal tolerability assumption and allowing for different AE rates between the treatments
- Including modelling the dose adjustment observed in the CELESTIAL and RESORCE trials
    - In the CELESTIAL trial, the mean daily dose of cabozantinib was 36.6 mg [46]. In the RESORCE trial the mean daily dose of regorafenib was 144.1 mg [41]. This scenario included a reduced cost for treatments based on the reduced number of whole packs needed to provide the total dosage received over the entire treatment duration. The total dosage received is calculated as the mean daily dose multiplied by treatment duration
- Assuming no drug wastage costs

The scenario results are shown in Table 40.

Table 40: Scenario analyses

Scenario	Difference in costs
Base case	[redacted]

Scenario	Difference in costs
Time on treatment – 20% ([redacted] months)	[redacted]
Time on treatment + 20% ([redacted] months)	[redacted]
Different toxicity profiles between treatments	[redacted]
Dose adjustments included	[redacted]
No drug wastage costs	[redacted]

Abbreviations: PFS, progression-free survival

B.4.5 Subgroup analysis

No subgroup analyses were considered as part of the cost-comparison.

B.4.6 Interpretation and conclusions of economic evidence

The cost-comparison analysis demonstrates that, when equivalent clinical effectiveness is assumed, cabozantinib is cost-saving when compared to regorafenib using the cabozantinib PAS price.

Regorafenib was selected as the comparator for the cost-comparison analysis because regorafenib is the only treatment option for patients that have received sorafenib in the preceding line of therapy and the positioning of cabozantinib in the treatment pathway is the same as regorafenib (Figure 2). The ITC showed that cabozantinib has similar or greater efficacy to regorafenib and as this analysis has demonstrated, cabozantinib is cost-saving in relation to regorafenib, which further supports the choice of the cost-comparison method.

In the analysis, only relevant costs, those associated with drug acquisition, were included as cabozantinib is not associated with any additional resource use as detailed above.

Scenario analyses all confirmed the base case analysis of cabozantinib as a cost-saving option. Increasing the time on treatment by 20% increased the base case cost savings by 34% and decreasing the time on treatment by 20% decreased the base case cost savings by 4%. When varying the time on treatment, the cost of drug wastage will impact the amount of savings with cabozantinib based on the required number of treatment cycles for each treatment. As the treatment duration increases, the number of cycles differs for each treatment since treatment cycle length is smaller for regorafenib. Consequently, an additional pack of regorafenib is costed compared

to the number of cabozantinib packs required. When relaxing the equal efficacy assumption, the costs from AEs were very small. Assuming additional cost savings from using a reduced dose increased the cost savings by 26%. The scenario with no drug wastage showed a 6% increase in savings. Therefore, cabozantinib offers the NHS an equally efficacious, cost-saving and tolerable alternative to regorafenib treatment of adults with advanced HCC who have previously been treated with sorafenib.

References

1. National Institute for Health and Care Excellence. *Guide to the processes of technology appraisal*. 2018 [cited 2022 January]; Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/technology-appraisal-processes-guide-apr-2018.pdf.
2. National Institute for Health and Care Excellence, *Regorafenib for previously treated advanced hepatocellular carcinoma TA555.* 2019.
3. National Institute for Health and Care Excellence, *Atezolizumab with bevacizumab for treating advanced or unresectable hepatocellular carcinoma TA666.* 2020.
4. Ipsen, Advisory board report, data on file. 2021.
5. National Health Service England. *National Cancer Drugs Fund List. ver1.200*. December 2021 [cited 2022 January]; Available from: https://www.england.nhs.uk/wp-content/uploads/2017/04/NationalCDF_List_ver1.200_-22122021_MASTER.pdf.
6. National Institute for Health and Care Excellence. *Single Technology Appraisal (STA). Cabozantinib for treating advanced hepatocellular carcinoma after prior therapy ID1243. Response to consultee and commentator comments on the draft remit and draft scope (pre-referral)*. 2018 [cited 2022 January]; Available from: https://www.nice.org.uk/guidance/ta582/documents/scope-consultation-comments-and-responses.
7. Abou-Alfa, G.K., et al., *Cabozantinib in Patients with Advanced and Progressing Hepatocellular Carcinoma.* N Engl J Med, 2018. **379**(1): p. 54-63.
8. Xiang Q, C.W., Ren M, Wang J, Zhang H, Deng DY, et al., *Cabozantinib suppresses tumor growth and metastasis in hepatocellular carcinoma by a dual blockade of VEGFR2 and MET.* Clinical cancer research : an official journal of the American Association for Cancer Research., 2014. **20**(11): p. 2959-70.
9. Yakes FM, C.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.* Molecular cancer therapeutics, 2011. **10**(12): p. 2298-308.
10. Bentzien, F., et al., *In vitro and in vivo activity of cabozantinib (XL184), an inhibitor of RET, MET, and VEGFR2, in a model of medullary thyroid cancer.* Thyroid : official journal of the American Thyroid Association, 2013. **23**(12): p. 1569-1577.
11. Escudier, B., J.C. Lougheed, and L. Albiges, *Cabozantinib for the treatment of renal cell carcinoma.* Expert Opin Pharmacother, 2016. **17**(18): p. 2499-2504.
12. Kurzrock, R., et al., *Activity of XL184 (Cabozantinib), an oral tyrosine kinase inhibitor, in patients with medullary thyroid cancer.* J Clin Oncol, 2011. **29**(19): p. 2660-6.
13. Kelley, R.K., et al., *Cabozantinib in hepatocellular carcinoma: results of a phase 2 placebo-controlled randomized discontinuation study.* Ann Oncol, 2017. **28**(3): p. 528-534.

14. Joint Formulary Committee. *British National Formulary (online) London: BMJ Group and Pharmaceutical Press <http://www.medicinescomplete.com> [Accessed on [01/12/2021]]*. 2022.
15. Orphanet., *Hepatocellular carcinoma. [15 June 2018]; Available from: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=EN&Expert=88673.*
16. Wu, J., *The changing epidemiology of hepatocellular carcinoma in Asia versus United States and Europe. .* Adv Mod Oncol Res, 2017. **3**: p. p. 51-58.
17. Blachier, M., et al., *The burden of liver disease in Europe: a review of available epidemiological data.* J Hepatol, 2013. **58**(3): p. 593-608.
18. El-Serag, H.B., *Epidemiology of viral hepatitis and hepatocellular carcinoma.* Gastroenterology, 2012. **142**(6): p. 1264-1273.e1.
19. Jemal, A., et al., *Global cancer statistics.* CA Cancer J Clin, 2011. **61**(2): p. 69-90.
20. McGlynn, K.A., J.L. Petrick, and W.T. London, *Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability.* Clin Liver Dis, 2015. **19**(2): p. 223-38.
21. Velázquez, R.F., et al., *Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis.* Hepatology, 2003. **37**(3): p. 520-7.
22. Burton, A., et al., *Primary liver cancer in the UK: Incidence, incidence-based mortality, and survival by subtype, sex, and nation.* JHEP Rep, 2021. **3**(2): p. 100232.
23. Cancer Research UK. *Liver Cancer incidence statistics.* [cited 2022 January]; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/incidence#heading-Zero.
24. Cancer Research UK. *Liver cancer mortality statistics.* [cited 2022 January]; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/mortality#heading-Zero.
25. EASL, *EORTC clinical practice guidelines: management of hepatocellular carcinoma.* J Hepatol, 2012. **56**(4): p. 908-43.
26. Vogel, A., et al., *Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2018. **29**(Suppl 4): p. iv238-iv255.
27. Bruix, J. and M. Sherman, *Management of hepatocellular carcinoma: an update.* Hepatology, 2011. **53**(3): p. 1020-2.
28. Durand, F. and D. Valla, *Assessment of prognosis of cirrhosis.* Semin Liver Dis, 2008. **28**(1): p. 110-22.
29. EASL, *Clinical Practice Guidelines: Management of hepatocellular carcinoma.* J Hepatol, 2018. **69**(1): p. 182-236.
30. Shiina, S., et al., *Percutaneous ethanol injection for hepatocellular carcinoma: 20-year outcome and prognostic factors.* Liver Int, 2012. **32**(9): p. 1434-42.
31. Shah, S.A., et al., *Recurrence after liver resection for hepatocellular carcinoma: risk factors, treatment, and outcomes.* Surgery, 2007. **141**(3): p. 330-9.
32. Bruix, J., M. Reig, and M. Sherman, *Evidence-Based Diagnosis, Staging, and Treatment of Patients With Hepatocellular Carcinoma.* Gastroenterology, 2016. **150**(4): p. 835-853.
33. Park, J.W., et al., *Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study.* Liver Int, 2015. **35**(9): p. 2155-66.

34. Verslype, C., O. Rosmorduc, and P. Rougier, *Hepatocellular carcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2012. **23 Suppl 7**: p. vii41-8.
35. Ryder, S.D., *British Society of Gastroenterology: Guidelines for the diagnosis and treatment of hepatocellular carcinoma (HCC) in adults.* Gut, 2003. **52**: p. iii1–iii8.
36. ESMO, *Hepatocellular carcinoma: European Society for Medical Oncology (ESMO) Clinical Practice Guidelines for diagnosis, treatment and follow-up (March 2021 update): .* 2021.
37. Bruix, J., et al., *Systemic treatment of hepatocellular carcinoma: An EASL position paper.* J Hepatol, 2021. **75**(4): p. 960-974.
38. *ILCA Systemic Therapy Guidance*. 2020 [cited 2022 January]; Available from: https://livercentral.org/wp-content/uploads/2021/12/Systemic-therapy-guidelines-V1.2.pdf.
39. National Institute for Health and Care Excellence, *Sorafenib for treating advanced hepatocellular carcinoma TA474.* 2017.
40. National Institute for Health and Care Excellence, *Lenvatinib for untreated advanced hepatocellular carcinoma TA551.* 2018.
41. Bruix, J., et al., *Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial.* Lancet, 2017. **389**(10064): p. 56-66.
42. Pinato, D.J., et al., *Integrated analysis of multiple receptor tyrosine kinases identifies Axl as a therapeutic target and mediator of resistance to sorafenib in hepatocellular carcinoma.* British Journal of Cancer, 2019. **120**(5): p. 512-521.
43. Xiang, Q.-F., et al., *Activation of MET promotes resistance to sorafenib in hepatocellular carcinoma cells via the AKT/ERK1/2-EGR1 pathway.* Artificial Cells, Nanomedicine, and Biotechnology, 2019. **47**(1): p. 83-89.
44. Firtina Karagonlar, Z., et al., *Elevated hepatocyte growth factor expression as an autocrine c-Met activation mechanism in acquired resistance to sorafenib in hepatocellular carcinoma cells.* Cancer Science, 2016. **107**(4): p. 407-416.
45. Llovet, J.M., et al., *Design and endpoints of clinical trials in hepatocellular carcinoma.* J Natl Cancer Inst, 2008. **100**(10): p. 698-711.
46. Exelixis, *CLINICAL STUDY REPORT XL184-309: A phase 3, randomized, double-blind, controlled study of cabozantinib (XL184) vs placebo in subjects with hepatocellular carcinoma who have received prior sorafenib*. 2018.
47. King, J., et al., *Sorafenib for the Treatment of Advanced Hepatocellular Cancer - a UK Audit.* Clin Oncol (R Coll Radiol), 2017. **29**(4): p. 256-262.
48. Llovet, J.M., et al., *Brivanib in patients with advanced hepatocellular carcinoma who were intolerant to sorafenib or for whom sorafenib failed: results from the randomized phase III BRISK-PS study.* J Clin Oncol, 2013. **31**(28): p. 3509-16.
49. Bucher HC, G.G., Griffith LE, Walter SD, *The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials.* J Clin Epidemiol 1997. **50(6):683-91.**: p. 683-91.
50. Finn, R.S., et al., *Outcomes of sequential treatment with sorafenib followed by regorafenib for HCC: Additional analyses from the phase III RESORCE trial.* J Hepatol, 2018. **69**(2): p. 353-358.
51. Guyot, P., et al., *Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves.* BMC Medical Research Methodology, 2012. **12**(1): p. 9.

52. Signorovitch, J.E., et al., *Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept.* Pharmacoeconomics, 2010. **28**(10): p. 935-45.
53. Signorovitch, J.E., et al., *Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research.* Value Health, 2012. **15**(6): p. 940-7.
54. Nash, P., et al., *Secukinumab Versus Adalimumab for Psoriatic Arthritis: Comparative Effectiveness up to 48 Weeks Using a Matching-Adjusted Indirect Comparison.* Rheumatology and therapy, 2018. **5**(1): p. 99-122.
55. Ipsen, Advisory board report, data on file. 2018.
56. Venables, W.N.a.R., B.D., *Modern Applied Statistics with S. Fourth Edition*. 2002.
57. White, H., *A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.* Econometrica, 1980. **48**(4): p. 817-838.
58. Woodruff, R.S., *Confidence Intervals for Medians and Other Position Measures.* Journal of the American Statistical Association, 1952. **47**(260): p. 635-646.
59. Mandel, M., *Simulation-Based Confidence Intervals for Functions With Complicated Derivatives.* The American Statistician, 2013. **67**(2): p. 76-81.
60. Kelley, R.K., et al., *Comparative Efficacy of Cabozantinib and Regorafenib for Advanced Hepatocellular Carcinoma.* Adv Ther, 2020. **37**(6): p. 2678-2695.
61. Casadei-Gardini, A., et al., *Regorafenib versus cabozantinb as second-line treatment after sorafenib for unresectable hepatocellular carcinoma: matching-adjusted indirect comparison analysis.* J Cancer Res Clin Oncol, 2021. **147**(12): p. 3665-3671.
62. Wang, D., et al., *Comparing the efficacy and safety of second-line therapies for advanced hepatocellular carcinoma: a network meta-analysis of phase III trials.* Therap Adv Gastroenterol, 2020. **13**: p. 1756284820932483.
63. Bakouny, Z., et al., *Second-line Treatments of Advanced Hepatocellular Carcinoma: Systematic Review and Network Meta-Analysis of Randomized Controlled Trials.* J Clin Gastroenterol, 2019. **53**(4): p. 251-261.
64. Sonbol, M.B., et al., *Systemic Therapy and Sequencing Options in Advanced Hepatocellular Carcinoma: A Systematic Review and Network Meta-analysis.* JAMA Oncology, 2020. **6**(12): p. e204930-e204930.
65. Park, R., et al., *Comparison of Efficacy of Systemic Therapies in Advanced Hepatocellular Carcinoma: Updated Systematic Review and Frequentist Network Meta-Analysis of Randomized Controlled Trials.* J Hepatocell Carcinoma, 2021. **8**: p. 145-154.
66. Tovoli, F., et al., *Real-Life Clinical Data of Cabozantinib for Unresectable Hepatocellular Carcinoma.* Liver Cancer, 2021. **10**(4): p. 370-379.
67. Finkelmeier, F., et al., *Cabozantinib in Advanced Hepatocellular Carcinoma: Efficacy and Safety Data from an International Multicenter Real-Life Cohort.* Liver Cancer, 2021. **10**(4): p. 360-369.
68. National Institute for Health and Care Excellence, *User guide for company evidence submission template: Fast track appraisal: cost-comparison case.* 2018.
69. Latimer, N., *NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL*

TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA in *REPORT BY THE DECISION SUPPORT UNIT*. 2011.
70. Ipsen, Parametric survival analyses of CELESTIAL ITT, data on file. 2021.
71. Ipsen, Clinical expert interviews. 2022.
72. National Health Service, *Reference costs 2019/20.* 2021.
73. Personal Social Service Unit, *Unit costs of health and social care.* 2021.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Document B

Company evidence submission

August 2022

File name	Version	Contains confidential information	Date
ID3917_Cabozantinib_HCC_Company Evidence Submission_Doc B_Final_STA_04.08.2022	**V4.0**	**Yes**	**04/08/2022**

B.1. Table of contents

List of Tables

List of Figures

Abbreviations

AEs	Adverse events
AIDS	Acquired immunodeficiency syndrome
AFP	Alpha-fetoprotein
ALT	Alanine aminotransferase
ANC	Absolute neutrophil count
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
BSC	Best supportive care
BOR	Best overall response
BP	Blood pressure
CCA	Cost-comparison analysis
CEAC	Cost-effectiveness acceptability curve
CEM	Cost-effectiveness model
CMH	Cochran-Mantel-Haenszel
CR	Complete response
CRF	Case report form
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
DSA	Deterministic sensitivity analysis
EASL	European Association for the Study of the Liver (EASL)
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D-5L	EuroQol 5-dimension, 5-level health questionnaire (EQ-5D-5L)
EQ-VAS	EQ-visual analogue scale
FTA	Fast track appraisal
GI	Gastrointestinal
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HR	Hazard ratio
HRQoL	Health-related quality of life
IDMC	Independent Data Monitoring Committee
ITT	Intention to treat
IxRS	Interactive voice/web response system (for randomisation)
MRI	Magnetic resonance imaging

NASH	Non-alcoholic steatohepatitis
NCCN	National Comprehensive Cancer Network
NHS	National Health Service
NMB	Net monetary benefit
NPACT	Non-protocol anticancer therapy
ORR	Objective response rate
OS	Overall survival
PD	Progressive disease
PD-1	Programmed cell death-1
PD-L1	Programmed cell death ligand 1
PFS	Progression-free survival
PPES	Palmar-plantar erythrodysaesthesia syndrome
PR	Partial response
PS	Performance status
QALY	Quality-adjusted life-year
QTcF	QT interval calculated by the Fridericia formula
RCT	Randomised controlled trial
RECIST	Response Evaluation Criteria In Solid Tumours
SAE	Serious adverse events
SD	Stable disease
TACE	Transarterial chemoembolisation
TTD	Time to treatment discontinuation
ULN	Upper limit of normal
UPCR	Urine protein/creatinine ratio
VEGF-A	Vascular endothelial growth factor A

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

B.1.1. Decision problem

The final marketing authorisation is:

Cabozantinib is indicated as monotherapy for the treatment of hepatocellular carcinoma (HCC) in adults who have previously been treated with sorafenib.

The population defined in the final scope is adults with advanced HCC who have had sorafenib.

The population defined in the final scope is adults with advanced HCC who have had sorafenib. The submission covers the technology's full marketing authorisation for this indication.

Comparator

The company submission differs from the final NICE scope with regards to the comparators. For cabozantinib the relevant comparator is regorafenib, as it is the only technology recommended in published NICE guidance for the same indication as cabozantinib. The wording of the regorafenib marketing authorisation is: *"Regorafenib is indicated as monotherapy for the treatment of adult patients with HCC who have been previously treated with sorafenib."*

Regorafenib is recommended by NICE (TA555) (1) as an option for treating advanced unresectable HCC in adults who have had sorafenib, only if:

> they have Child Pugh grade A liver impairment and an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1.
> the company provides it according to the commercial arrangement.

The NICE recommendation includes a restriction on the eligible patient population based on the degree of liver impairment and performance status. This is because the clinical trial evidence for regorafenib is based on advanced HCC patients that have been previously treated with sorafenib, and who have an ECOG performance status

of 0 or 1 and Child-Pugh grade A liver impairment and not those who have more severe liver disease or a poorer performance status. Following the NICE approval of atezolizumab plus bevacizumab in first line (TA666) (2), regorafenib is now also being used in the third-line setting (further details regarding the treatment pathway are found in Section B.1.3). The positioning and use of regorafenib as the comparator for cabozantinib in clinical practice has been validated by clinical experts (3) treating eligible patients with drugs that are reimbursed according to the National Health Service England (NHSE) National Cancer Drugs Fund List (NCDFL) (4).

Ipsen wish to pursue the same positioning as the NICE recommendation for regorafenib as the clinical trial evidence is relatively limited for cabozantinib in people with advanced HCC with more severe liver disease or a poorer performance status.

It should be noted that best supportive care (BSC) is not a relevant comparator for cabozantinib, as the comparator can only be technologies already recommended in published technology appraisal guidance and/or treatment guidelines for the same indication.

Several analyses were conducted to provide evidence to support the comparative effectiveness of cabozantinib versus regorafenib, which consistently support similar or greater efficacy of cabozantinib versus regorafenib. A summary of the evidence includes the following:

- Indirect treatment comparisons (ITCs) using well-accepted and validated methodologies were conducted. The findings show no clear trend in ITC results in favour of cabozantinib versus regorafenib, but it can be concluded that cabozantinib is at least similar in clinical effectiveness to regorafenib and this conclusion is further supported from real world evidence (RWE) findings. The results of these analyses and RWE are described in greater detail in Section 3.10;
- Cabozantinib and regorafenib belong to same drug class of tyrosine kinase inhibitors (TKIs). They inhibit multiple receptor tyrosine kinases (RTKs) implicated in tumour growth, metastasis, and angiogenesis, including vascular endothelial growth factor receptor (VEGFR), endothelial-specific Angiopoietin receptor (TIE-2), mast/stem cell growth factor (KIT) and rearranged during

transfection receptor (RET). The safety profile of cabozantinib is generally similar to that of other VEGFR-targeting TKIs. The results of adverse event comparisons are described in greater detail in Section 3.10;

- Clinical experts were consulted in an advisory board conducted by the manufacturer (3). The clinical experts believe that the clinical effectiveness of cabozantinib and regorafenib are broadly equivalent. This is also reflected in the responses from the British Association for the Study of the Liver (BASL)/HCC-UK, British Society of Gastroenterology (BSG) and the National Cancer Research Institute (NCRI) Hepatobiliary Working Group, to the NICE scoping consultation for this topic; the NCRI Hepatobiliary Working Group also felt the FTA cost-comparison route was also appropriate (5).

To fulfil the criteria of “similar or lower costs”, [redacted]. As previously mentioned, regorafenib is the only approved NICE therapy in this indication, which is further supported by the findings of the clinical expert advisory board (3).

The decision problem addressed by this submission is summarised in Table 1.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with advanced hepatocellular carcinoma who have had sorafenib.	Adults with advanced hepatocellular carcinoma who have had sorafenib.	N/A
Comparator(s)	Regorafenib Best supportive care (BSC)	Regorafenib	BSC is not a relevant comparator in a cost-comparison case as the comparator can only be technologies already recommended in published technology appraisal guidance and/or treatment guidelines for the same indication.
Outcomes	• Overall survival • Progression-free survival • Response rates • Time to treatment discontinuation • Adverse effects of treatment • Health-related quality of life	• Overall survival • Progression-free survival • Response rates • Time to treatment discontinuation • Adverse effects of treatment • Health-related quality of life	The published literature for regorafenib does not present time to treatment discontinuation, limiting a comparison using this outcome
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. If the technology is likely to provide similar or greater health benefits at similar or lower cost than technologies recommended in published NICE technology appraisal guidance for the same indication, a cost-comparison may be carried out. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or	As per final scope.	N/A

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	outcomes between the technologies being compared. Costs will be considered from a NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention, comparator or subsequent treatment technologies will be taken into account.		
Subgroups to be considered	None specified.	None specified.	N/A
Special considerations including issues related to equity or equality	N/A	No equity or equality issues for consideration	N/A

Abbreviations: NHS, National Health Service; NICE, The National Institute for Health and Care Excellence.

B.1.2. Description of the technology being appraised

This appraisal considers the proposed indication for cabozantinib patients with hepatocellular carcinoma (HCC) who have been previously treated with sorafenib. Cabozantinib's mechanism of action, marketing authorisation, indication, mode of administration and list price are summarised in Table 2.

Please refer to Appendix C which includes the draft summary of product characteristics (SmPC) for this technology, pending finalisation of the marketing authorisation process.

Table 2: Technology being appraised

UK approved name and brand name	Cabozantinib (Cabometyx®)
Mechanism of action	Cabozantinib is an oral multi-targeted inhibitor of receptor tyrosine kinases (RTKs) that potently inhibits several RTKs known to influence tumour growth, metastasis and angiogenesis, including MET, VEGFR2 and AXL (6) . Treatment with cabozantinib results in anti-angiogenic effects in xenograft tumours, with disruption of the vasculature beginning within 24 hours after administration and is associated with pro-apoptotic effects leading to significant tumour growth inhibition or tumour regression in multiple tumour models including HCC, medullary thyroid cancer (MTC), breast cancer, lung carcinoma, glioblastoma and renal cell carcinoma (7-10). The broad clinical activity of cabozantinib was demonstrated in a Phase I trial, in which tumour regression was observed in multiple tumour types (11) and these early findings were confirmed in a phase II randomised discontinuation trial (XL184-203 RDT) conducted in 9 tumour types, including HCC (12).
Marketing authorisation/CE mark status	An application for the marketing authorisation for cabozantinib in this indication was submitted to the European Medicines Agency (EMA) on 31 March 2018. The marketing authorisation process for the United Kingdom (UK) was centralised through the EMA at that time. The EMA granted marketing authorisation for cabozantinib, as monotherapy for the treatment of HCC in adults who have previously been treated with sorafenib, in November 2018.
Indications and any restriction(s) as described in the summary of	Cabozantinib is indicated as monotherapy for the treatment of HCC in adults who have previously been treated with sorafenib.

product characteristics	See Appendix C for the Summary of Product Characteristics and European public assessment report (EPAR).
Method of administration and dosage	Oral administration: One 60mg tablet to be taken once daily. Management of suspected adverse drug reactions may require temporary treatment interruption and/or dose 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. If a patient misses a dose, the missed dose should not be taken if it is less than 12 hours before the next dose.
Additional tests or investigations	A biopsy is required to establish histological or cytological diagnosis of HCC. Radiographic tumour assessment was performed every eight weeks using computed tomography or magnetic resonance imaging to assess disease progression in the pivotal trial. It is also recommended to monitor biochemical and metabolic parameters during treatment. This monitoring would likely be carried out as part of the routine management of advanced HCC.
List price and average cost of a course of treatment	£5,143.00 per 30 tablet pack. (13) Average cost per course of treatment equal to to [redacted]
Patient access scheme (if applicable)	A confidential simple patient access scheme is available. The pack price under this scheme is [redacted] (a [redacted]% discount to the list price). Under this scheme the average cost of a course of treatment (based on a [redacted]) is £[redacted].

Abbreviations: EMA, European Medicines Agency; EPAR, European public assessment report; HCC, hepatocellular carcinoma; MET, mesenchymal epithelial transition factor; MTC, medullary thyroid cancer; RDI, relative dose intensity; RTKs, receptor tyrosine kinases, UK: United Kingdom; VEGFR, vascular endothelial growth factor receptor.

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

B.1.3.1 Disease overview

HCC is a primary hepatic cancer derived from well-differentiated hepatocytes (14). It is the most common histologic subtype of liver cancer (15), accounting for approximately 80% of all liver cancers cases (estimates range from 70 to 90%) (16-20). HCC occurs predominantly in patients with underlying chronic liver disease and cirrhosis; 70–90% of HCC cases develop against a background of cirrhosis, typically

associated with hepatitis liver (chronic hepatitis B and/or hepatitis C virus infection), alcohol consumption, non-alcoholic steatohepatitis and haemochromatosis (16).

In the United Kingdom (UK), HCC is amongst cancers with the most rapid rate of growth both in incidence and mortality in last few decades (21). There are 6,214 new cases of liver cancer each year in UK (2016-2018) with 66% cases in males (22). The European Age-Standardised incidence rate in the UK (2016-2018) was 10 per 100,000 population; with significantly higher rates in men (14.5 per 100,000) compared to women (6.2 per 100,000) (22). Over the last decade (between 2006-2008 and 2016-2018), the liver cancer age-standardised incidence rate increased by 45% in the UK (22). It is projected to rise by 38% between 2014 and 2035, to 15 per 100,000 people by 2035 (22). Approximately 5,600 deaths are caused by liver cancer in the UK every year, accounting for 3% of all cancer deaths in 2018 (23). The liver cancer age-standardised mortality rates increased by 48% in the UK over the last decade (23), which is projected to rise by 58% between 2014 and 2035, to 16 deaths per 100,000 people by 2035 (23).

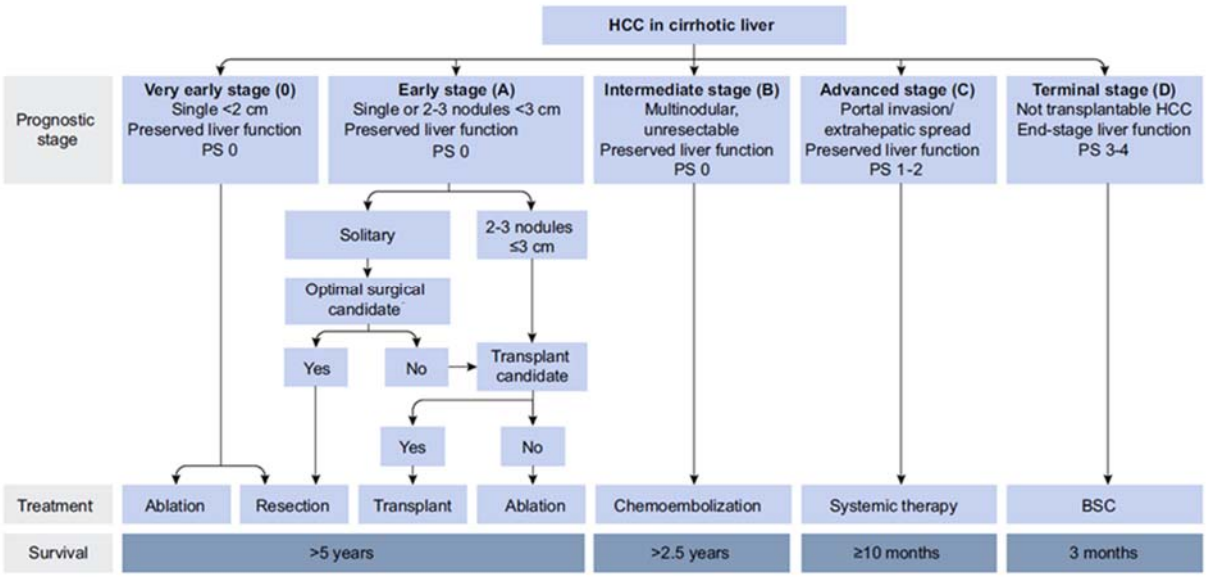
The overall prognosis for HCC depends on the severity of underlying liver dysfunction at the time of diagnosis as defined by the disease stage; the prognosis remains poor due to rapid disease progression and low survival rates. The age-standardised net survival rate at 1 year is 38.1%, and the net survival rate at 5 years is 12.7% for liver cancer, in England (23).

There are numerous disease staging systems for HCC, of which the Barcelona Clinic Liver Cancer (BCLC) staging system is most widely used in the UK. The BCLC has received the endorsements of the European Association for the Study of Liver Diseases (EASL) (24), the European Society for Medical Oncology (ESMO) (25) and the American Association for the Study of Liver Diseases (AASLD) (26). The BCLC classification divides HCC patients into five stages (0, A, B, C and D) considering prognostic variables related to tumour status, liver function (as measured by the Child–Pugh score) and health performance status (as measured by Eastern Cooperative Oncology Group [ECOG]), along with treatment-dependent variables identified from cohort studies and randomised trials. The Child-Pugh status, which measures the severity of cirrhosis, takes into account five clinical measures and scores them

between 1 and 11. This leads to a classification of Child-Pugh A, B or C, with C being the most severe (27). The classification of HCC is illustrated in Figure 1 and outlined in

Table 3.

Figure 1: Barcelona Clinic Liver Cancer (BCLC) staging system and treatment strategy (EASL Guidelines)



Source: EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma (28)

Table 3: Staging of HCC using BCLC classification

BCLC Staging	Tumour status	ECOG performance status	Liver function (Child-Pugh)
Stage 0 (Very early HCC)	Singe tumour <2cm in diameter without vascular invasion/satellites	0	Well preserved function Child-Pugh A
Stage A (Early HCC)	Single tumours >2cm or up to 3 nodules <3 cm in diameter	0	Child-Pugh A or B
Stage B (Intermediate HCC)	Multinodular asymptomatic tumours without an invasive pattern	0	Child-Pugh A or B
Stage C (Advanced HCC)	Symptomatic tumours; macrovascular invasion (either segmental or portal invasion) or extrahepatic spread (lymph node involvement or metastases)	0-2*	Child-Pugh A or B
Stage D (End stage HCC)		3-4	Child-Pugh C

*ESMO guidelines describe Stage C (advanced HCC) with ECOG performance status of 0-2 (25)
Abbreviations: BCLC staging, Barcelona clinic liver cancer staging; ECOG, Eastern Cooperative Oncology Group; HCC, Hepatocellular carcinoma.

Treatment for HCC depends on the location and stage of the cancer, and how well the liver function is preserved. Approximately 30-40% of HCC patients worldwide, who are diagnosed with very early or early disease (BCLC stage 0/A), are eligible for curative procedures, which may include surgery (hepatic resection or liver transplantation) or percutaneous ablation (26, 28-30). Around half of patients with HCC undergoing resection have a relapse in less than three years (30).

Patients with intermediate-stage HCC (BCLC B), in whom liver function is preserved, may be candidates for transarterial chemo-embolisation (TACE) (31). Most patients are diagnosed in the advanced stages of the disease (BCLC stage C), when cirrhosis is present, when surgery is rarely an option and treatment is palliative rather than curative (32). Half of those diagnosed with BCLC stage C do not survive for more than 3 months. Without treatment, the median survival for BCLC stage C patients ranges between 4 and 8 months (33).

B.1.3.2 Current treatment pathway for advanced HCC

For patients with advanced HCC, treatment options include interventional procedures such as TACE (using doxorubicin or cisplatin) or selective internal radiation therapy, and external beam radiotherapy. Patients who do not respond to these therapies, or have metastatic disease, are treated with systemic therapies.

UK and European clinical practice guideline recommendation

The British Society of Gastroenterology (BSG) guidelines for HCC in UK clinical practice, were published in 2003, prior to sorafenib and regorafenib becoming available (34). As this existing guideline is outdated, the UK clinical practice largely aligns with the NICE treatment pathway and European guidelines published by ESMO and EASL (25, 35). The EASL guidelines, published in 2018 before cabozantinib had received EMA's approval in HCC, recommend sorafenib and lenvatinib in first-line and regorafenib in second-line (25). The 2021 EASL position paper complements the 2018 guidance, recommending sorafenib, atezolizumab plus bevacizumab, and lenvatinib as first-line treatments (36). In second-line, post atezolizumab plus bevacizumab, the 2021 EASL position paper recommends multi targeted tyrosine kinase inhibitors (TKIs) and vascular endothelial growth factor receptor-2 (VEGF-2) TKIs (36). In second-line

post sorafenib and lenvatinib, regorafenib, cabozantinib and ramucirumab are recommended treatments (36).

The ESMO clinical practice guidelines (35), updated more recently in March 2021, recommends atezolizumab plus bevacizumab, sorafenib and lenvatinib as first-line treatments. After treatment with sorafenib, the guidelines recommend cabozantinib, regorafenib and ramucirumab as 'standard' second-line treatments. As there is no evidence for any drug in particular, ESMO guidelines recommend that all the currently approved first- and second-line agents could be considered as second-line therapy post atezolizumab plus bevacizumab i.e. sorafenib, lenvatinib, cabozantinib, regorafenib and ramucirumab (35). It is noteworthy that regorafenib is not recommended for TKI-naive patients by the ESMO guidelines, after treatment with either sorafenib or atezolizumab plus bevacizumab.

The 2020 International Liver Cancer Association (ILCA) guidelines also recommend atezolizumab plus bevacizumab as standard of care, with exception in patients for whom atezolizumab or bevacizumab are contraindicated (sorafenib and lenvatinib are recommended as alternative option) (37). Although, there is no data to support one TKI over another, the ILCA guidelines suggest sorafenib, lenvatinib and cabozantinib, after treatment with atezolizumab plus bevacizumab in first-line (37). The ILCA guidelines also supported the use of regorafenib (in patients who tolerated sorafenib) and ramucirumab (in patients with alpha fetoprotein (AFP) ≥400 ng/mL) (37). The ILCA guidelines also suggest sorafenib, cabozantinib, regorafenib, and ramucirumab as options if lenvatinib is used first-line, although there are no data to support this (37).

Atezolizumab plus bevacizumab is the first treatment to demonstrate a significant OS benefit compared with sorafenib and consequently the treatment landscape has changed with atezolizumab plus bevacizumab becoming the standard of care in first-line systemic therapy for advanced HCC in the UK (3). For patients having treatment with atezolizumab plus bevacizumab, the median progression-free survival (PFS) is only 6.8 months, raising the need to define options for second-line therapy (35, 37). Drugs in the second-line setting have so far only been tested after sorafenib failure/intolerance and there are currently no phase III trial data to inform the choice of second-line therapy in HCC patients that received alternative front-line therapies.

There is, however, a clear rationale for offering a multikinase inhibitor given the existing evidence for efficacy in first and second-line.

The EASL, EMSO and ILCA guidelines are summarised in Table 4.

Table 4: EASL, ESMO and ILCA summary of guidelines

	ESMO 2021	EASL 2021	ILCA 2020
First-line	Standard: • Atezolizumab plus bevacizumab Option: • Sorafenib • Lenvatinib	Atezolizumab plus bevacizumab If contraindications to atezolizumab plus bevacizumab: • Sorafenib • Lenvatinib	First choice: Atezolizumab plus bevacizumab Alternative: • Sorafenib • Lenvatinib
Second-line	Option post atezolizumab plus bevacizumab: Cabozantinib Sorafenib Lenvatinib Regorafenib (only in patients previous exposed to TKIs) Ramucirumab (only in patients with an AFP level ≥400 ng/mL) Standard post-sorafenib: • Cabozantinib Regorafenib (only in patients previous exposed to TKIs) Ramucirumab (only in patients with an AFP level ≥400 ng/mL)	Post atezolizumab plus bevacizumab: Multi-TKI and VEGFR2 inhibitor as per off-label availability Post-sorafenib or lenvatinib: • Cabozantinib Regorafenib (in sorafenib-tolerant patients) • Ramucirumab (in patients with serum alpha--fetoprotein above 400 ng/ml)	Post atezolizumab plus bevacizumab: Cabozantinib Sorafenib Lenvatinib Post-sorafenib: • Cabozantinib Regorafenib (in patients who tolerated sorafenib) • Ramucirumab (If AFP ≥400 ng/mL) Post-lenvatinib first line: • Sorafenib • Cabozantinib Ramucirumab (if AFP ≥400 ng/mL)

	ESMO 2021	EASL 2021	ILCA 2020
Third-line		• Regorafenib (in patients who tolerated sorafenib)	• Regorafenib (in patients who tolerated sorafenib)

Abbreviations: AFP, alpha fetoprotein; EASL, European Association for the Study of Liver Diseases; ESMO, European Society for Medical Oncology; ILCA, International Liver Cancer Association; TKI, tyrosine kinase inhibitor
Source: EASL 2021 (36), ESMO 2021 (35), ILCA 2020 (37)

NICE recommendations for first-line systemic treatment of advanced and unresectable HCC

NICE has recommended atezolizumab plus bevacizumab combination, sorafenib, and lenvatinib as first-line systemic therapies for adult patients with advanced and unresectable HCC (TA666 (2), TA474 (38), TA551 (39)). NICE guidance based on technology appraisals are presented in Table 5 and Figure 2 (3).

NICE recommendations for second and later-line systemic treatment of advanced and unresectable HCC

For patients who have had sorafenib, only one treatment option, regorafenib, currently exists and is the standard of care in the UK practice following treatment with sorafenib. NICE has recommended regorafenib for patients who have had sorafenib, only if they had Child–Pugh grade A liver impairment and an ECOG performance status of 0 or 1 (TA555 (1)). The clinical evidence for regorafenib, however, is based on the RESORCE trial, which studied regorafenib as a second-line treatment option for patients receiving and tolerating sorafenib in the first-line setting. NICE guidance based on this technology appraisal is presented in Table 5 and Figure 2 (3). Ramucirumab is not currently approved by NICE, and currently, there is no ongoing NICE appraisal for ramucirumab. UK clinical experts have also confirmed to Ipsen that ramucirumab is not used in UK clinical practice (3).

Table 5: Summary of NICE technology appraisals related to HCC

NICE Technology Appraisals	Date
Atezolizumab with bevacizumab for advanced or unresectable hepatocellular carcinoma (TA666) – *Atezolizumab plus bevacizumab is recommended as an option for treating advanced or unresectable hepatocellular carcinoma (HCC) in adults who have not had previous systemic treatment, only if they have Child-Pugh grade A liver impairment and an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1, and the company provides it according to the commercial arrangement.*	December 2020 NICE TA666

NICE Technology Appraisals	Date
Regorafenib for previously treated unresectable hepatocellular carcinoma (TA555) – *Regorafenib is recommended as an option for treating advanced unresectable hepatocellular carcinoma in adults who have had sorafenib, only if they have Child-Pugh grade A liver impairment and an ECOG performance status of 0 or 1, and the company provides it according to the commercial arrangement.*	January 2019 NICE TA555 (replaces TA514)
Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma (TA551) – *Lenvatinib is recommended as an option for untreated, advanced, unresectable hepatocellular carcinoma in adults, only if patients have Child-Pugh grade A liver impairment and an ECOG performance status of 0 or 1, and the company provides lenvatinib within the agreed commercial arrangement.*	December 2018 NICE TA551
Sorafenib for the treatment of advanced hepatocellular carcinoma (TA474) – *Sorafenib is recommended as an option for treating advanced hepatocellular carcinoma only for people with Child-Pugh grade A liver impairment, only if the company provides sorafenib within the agreed commercial access arrangement.*	September 2017 NICE TA474 (replaces TA189)

Abbreviations: ECOG, Eastern Cooperative Oncology Group; HCC, hepatocellular carcinoma; NICE, The National Institute for Health and Care Excellence; TA, technology appraisal.

NHS England National Cancer Drugs Fund List (NHSE NCDFL)

The actual position of NICE recommended medicines that are reimbursed by NHSE NCDFL (4) is slightly different to the wording of the NICE recommendations. For lenvatinib and sorafenib use in second line (despite all the evidence only being for their use in a first-line setting), there are additional specific criteria applied to account for the changing landscape, as atezolizumab plus bevacizumab becomes the standard of care in first-line systemic therapy in advanced HCC. These additional criteria, that are outside of NICE recommendations include:

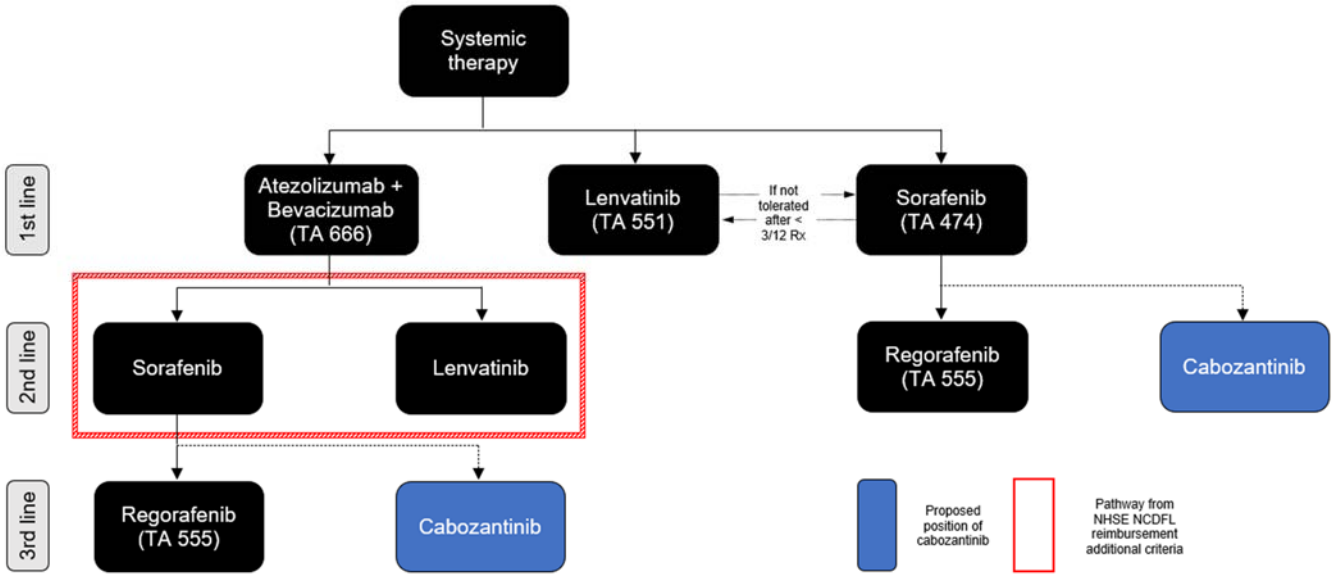
- Ability to receive lenvatinib or sorafenib if the patient has received atezolizumab plus bevacizumab as first-line treatment.
- Ability to switch from lenvatinib to sorafenib (and vice versa) in the first-line setting if patient has had to discontinue treatment within 3 months of starting the drug and solely because of toxicity.

Thus, regorafenib is currently prescribed and reimbursed post sorafenib by NHSE in either a second-line setting where sorafenib has been prescribed first-line or in a third-

line setting (despite the lack of evidence demonstrating its efficacy beyond the second-line treatment setting) for patients previously treated with atezolizumab plus bevacizumab followed by sorafenib. There is no NICE approved recommendation for second-line treatment following first-line treatment with lenvatinib, although clinical experts would welcome a treatment option in this setting (3).

The positioning of these treatments has been confirmed by UK clinical experts (3) and is summarised in Figure 2 below.

Figure 2: Current systemic therapy treatment pathway in UK clinical practice as per NICE and NHSE NCDFL recommendations



Abbreviations: NCDFL, National Cancer Drug Fund List; NHSE, National Health Service England; NICE, National Institute for Health and Care Excellence; Rx, prescription.
Source: Clinical experts' opinion (3), NHSE NCDFL (4).

B.1.3.3 Positioning of Cabozantinib

It is proposed that cabozantinib is positioned where regorafenib is currently used in practice as shown in Figure 2.

It can be argued that the evidence base and generalisability of cabozantinib for the UK advanced HCC population is greater than that of regorafenib in its current position for the following reasons:

- The pivotal clinical trial (CELESTIAL) (6) for cabozantinib had broader inclusion criteria than that of the regorafenib pivotal trial (RESORCE) (40) as it included:
    - Both second and third-line patients (28% of trial patients were receiving third-line therapy) whilst the RESORCE trial only included patients who had received sorafenib first-line only i.e. the regorafenib population were pure second-line;
    - Patients intolerant to sorafenib. The CELESTIAL trial included patients who had disease progression on sorafenib irrespective of whether they had tolerated sorafenib or not, unlike the RESORCE trial where patients who had disease progression on sorafenib had to have tolerated sorafenib (≥400 mg daily for at least 20 of the 28 days before discontinuation);
    - Additionally, compared to the RESORCE trial patients in the CELESTIAL trial were more likely to be white (56% versus 36%), and less likely to be in the Asia geographical region (25% versus 38%).

This makes cabozantinib a more relevant treatment option than regorafenib in practice when taking into account the current NICE and NHSE NCDFL recommendations.

Due to differences in the CELESTIAL and RESORCE trials designs, no superiority claim is made for cabozantinib in this submission. However, cabozantinib is currently the only therapy developed for HCC that inhibits the MET and AXL receptors (in addition to VEGFR 1, 2 and 3), and thereby provides additional inhibitory effects beyond that of currently approved TKIs (8). Due to this unique molecular pathway, cabozantinib may be able to break TKI resistance established in the first-line of treatment (41-43). Therefore, cabozantinib has a biologically plausible rationale to treat patients who are resistant to sorafenib (6).

Cabozantinib has demonstrated to be efficacious in a broader patient population of advanced HCC previously treated with sorafenib in the pivotal CELESTIAL trial, which was a robust, double-blind randomised trial investigating the impact of cabozantinib

compared with placebo (6). At the time of the design of the CELESTIAL trial, there were no other treatments available other than best supportive care (BSC), and, therefore, placebo was used as the comparator arm of the trial.

Cabozantinib demonstrated a statistically significant improvement in OS versus placebo from 8.0 months to 10.2 months. This amounts to a 24% reduction in risk of death (44) in this population, which is more representative of the real-world population than the clinical evidence from the RESORCE trial. The EASL guidelines state that cabozantinib has shown survival benefits vs. placebo in the second-line setting (25).

The proposed position of cabozantinib as a treatment option after prior treatment with sorafenib offers an alternative treatment option to a UK patient population with poor prognosis where there is only one other treatment option currently recommended by NICE. For these patients, cabozantinib offer an additional treatment option, including patients intolerant to sorafenib.

Similarly, cabozantinib's proposed position as a third-line treatment option after initial treatment with atezolizumab plus bevacizumab followed by sorafenib, would provide patients not only with an alternative treatment option other than regorafenib, but also serves as the only available treatment option with proven efficacy in a third line setting that is based on clinical trial evidence.

Access to cabozantinib for UK HCC patients provides not only the option but also the reassurance to both patients and providers that they are receiving a treatment option demonstrated to be efficacious for a broader patient population with advanced HCC (6).

This submission aims to demonstrate that cabozantinib is a health technology that is likely to provide similar or greater health benefits at similar or lower cost than technologies already recommended (i.e., regorafenib) in published technology appraisal guidance for the same indication. In addition, it aims to demonstrate that the evidence base for cabozantinib is more generalisable to UK practice and thus offers an additional treatment option for UK patients with advanced HCC, where systemic treatment options are limited and the prognosis remains poor as they continue to progress rapidly and have a short overall survival (OS) of 8 to 11 months (24, 40).

B.1.4. Equality considerations

No equality issues related to the use of cabozantinib have been identified.

B.2. Clinical effectiveness

Cabozantinib significantly improved OS, PFS and the ORR compared with placebo, with a manageable safety profile.

Clinical efficacy and Safety

- Cabozantinib significantly extended OS in advanced HCC patients versus placebo: median OS 10.2 months (95% CI: 9.1, 12.0) for cabozantinib versus 8.0 months (95% CI: 6.8, 9.4) for placebo, with a hazard ratio (HR) for death: 0.76 (95% CI: 0.63, 0.92; P = 0.005) (6)
- In the subgroup analysis of patients previously treated with sorafenib only, cabozantinib provided an additional 4.1 months of median OS versus placebo (11.3 months for cabozantinib and 7.2 months for placebo). Risk of death was reduced by 30% in this population (stratified HR for death: 0.70; 95% CI: 0.55, 0.88) (6)
- Cabozantinib significantly improved PFS in advanced HCC patients: median PFS 5.2 months (95% CI: 4.0, 5.5) versus 1.9 months (95% CI: 1.9, 1.9) for placebo, with a HR for disease progression or death: 0.44; (95% CI: 0.36, 0.52; P<0.001) (6)
- In the subgroup analysis of patients previously treated with sorafenib only, cabozantinib provided an additional 3.6 months of median PFS (5.5 months for cabozantinib and 1.9 months for placebo; HR for disease progression or death: 0.40; 95% CI: 0.32, 0.50) (6)
- AEs were consistent with the known safety profile of cabozantinib (6)

B.2.1. Identification and selection of relevant studies

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

B.2.2. List of relevant clinical effectiveness evidence

Clinical evidence to support the use of cabozantinib (XL184) for the treatment of advanced hepatocellular carcinoma (HCC) comprises a single randomised controlled trial (RCT) – the CELESTIAL trial (XL184-309; NCT01908426). A brief overview of this trial is provided in Table 6.

A systematic review of the literature did not identify any additional studies relevant to cabozantinib in advanced HCC.

Table 6: Clinical effectiveness evidence

Study	CELESTIAL
Study Design	Randomised, double-blind, placebo-controlled, phase III
Population	Patients with previously treated advanced HCC
Intervention(s)	Oral cabozantinib 60 mg once daily plus best supportive care (BSC)
Comparators	Oral matched placebo once daily plus BSC
Does trial support application for marketing authorisation	Yes
If trial 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; EQ-5D-5L, Health-related quality of life; HCC, hepatocellular carcinoma; 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

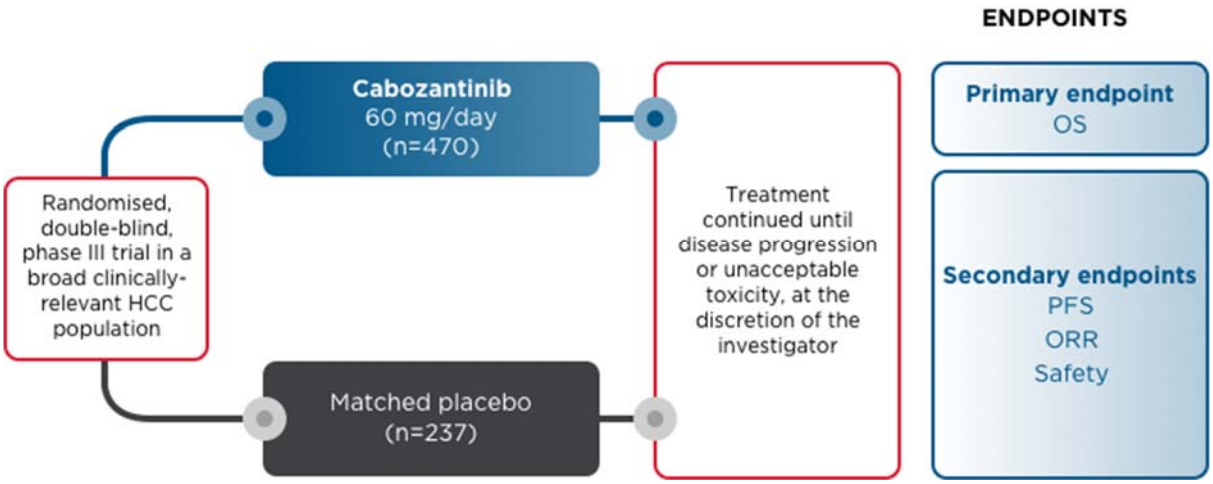
CELESTIAL Trial: The CELESTIAL global phase III clinical trial tested the effects of cabozantinib compared with placebo in patients with advanced HCC who had already received treatment with sorafenib.

B.2.3.1 Trial design

The CELESTIAL trial was a randomised, double-blind, placebo-controlled, phase III trial undertaken to assess the safety and efficacy of cabozantinib compared with placebo in patients with advanced HCC who had received prior treatment including sorafenib (6).

The schematic design of the trials is depicted in Figure 3.

Figure 3: CELESTIAL Trial design



Abbreviations: HCC, hepatocellular carcinoma; ORR, objective response rate; OS, overall survival; PFS, progression-free survival.
Source: Abou-Alfa et al., 2018 (6).

Outcome measures used in the economic model or specified in the scope

The relevant endpoints in the CELESTIAL trial along with details of when and how they were measured during the trial are summarised in Table 7 (44). All endpoints and outcomes described were pre-specified, unless otherwise stated.

Table 7: Relevant endpoints and measures in the CELESTIAL trial

Endpoint	Definition	Timing and nature of assessment
Primary endpoint		
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
Secondary endpoints		
PFS	The date of randomisation to radiographical progression or death, whichever occurred first	Radiographical tumour assessment by the investigator (or radiologist) was based on Response Evaluation Criteria in Solid Tumours (RECIST 1.1) Computed tomography (CT)/magnetic resonance imaging (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) Bone scans were performed at screening, 8 and 16 weeks after randomisation, and every 16 weeks in patients with documented bone lesions at screening or suspicion of bone metastasis during the trial 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
ORR	The proportion of patients with a best overall response (BOR) of complete response (CR) or partial response (PR). CR or PR must be confirmed on a subsequent visit ≥28 days after the response was first observed	
Exploratory endpoints		
HRQoL	Health status was measured using EQ-5D-5L 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	
Safety and tolerability	Safety assessments included the evaluation of AEs, serious AEs (SAEs), deaths, clinical laboratory tests (haematology, serum chemistry and urinalysis), physical examination, vital signs, ECOG PS, 12-lead electrocardiogram (ECG) and the TTD in months (date of decision to discontinue study drug – date of first dose +1)/30.4375.	

Endpoint	Definition	Timing and nature of assessment
	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.	

Abbreviations: AE, adverse event; ECOG, Eastern Cooperative Oncology Group; HRQoL, health-related quality of life; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PS, performance status; TTD, time to treatment discontinuation.

A summary of the methodology of the Phase III CELESTIAL trial is presented in Table 8.

Table 8: Summary of trial methodology: CELESTIAL trial

Study	CELESTIAL TRIAL
Location	104 sites across 19 countries (Australia, Belgium, Canada, France, Germany, Hong Kong, Ireland, Italy, Republic of Korea, Netherlands, New Zealand, Poland, Romania, Singapore, Spain, Taiwan, Turkey, UK and USA)
Trial Design	Phase III, randomised, double-blind, controlled study of Cabozantinib versus Placebo in patients with HCC who have received prior Sorafenib
Eligibility criteria for participants	Inclusion criteria included: • Age ≥18 years of age on the day of consent • Histological or cytological diagnosis of HCC (previous biopsy results accepted) • Disease not amenable to curative treatment (e.g., transplant, surgery, radiofrequency ablation) • Received prior sorafenib • Progression following at least one prior systemic treatment for HCC • Recovery from toxicities related to any prior treatment to ≤Grade 1, unless the AEs were clinically non-significant and/or stable with supportive therapy • Eastern Cooperative Oncology Group (ECOG) performance status (PS): 0 or 1 at screening • Adequate haematological function, i.e., meeting the following laboratory criteria ≤7 days prior to randomisation: • Absolute neutrophil count (ANC): ≥1200/mm3 (≥1.2×109/L) • Platelets: ≥60,000/mm3 (≥60×109/L) • Haemoglobin: ≥8 g/dL (≥80 g/L) • Adequate renal function, i.e., meeting the following laboratory criteria ≤7 days prior to randomisation: • Serum creatinine ≤1.5´upper limit of normal (ULN) or calculated creatinine clearance ≥40 mL/min using the Cockcroft-Gault equation • Urine protein/creatinine ratio (UPCR) ≤1 mg/mg (≤113.1 mg/mmol) or 24-hour urine protein <1 g • Child-Pugh status: A • Total bilirubin ≤2 mg/dL (≤34.2 µmol/L) ≤7 days prior to randomisation • Serum albumin ≥2.8 g/dL (≥28 g/L) ≤7 days prior to randomisation • Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) <5.0×ULN ≤7 days prior to randomisation • Haemoglobin A1c (HbA1c) ≤8% within 28 days prior to randomisation (if HbA1c results were unavailable: fasting serum glucose ≤160 mg/dL) • If have active HBV infection, receiving antiviral therapy according to the local standard of care • Be capable of understanding and complying with the protocol requirements and providing written consent • Sexually active fertile subjects and their partners must have agreed to use medically accepted barrier methods of contraception during the trial and for 4 months after the last dose of study drug

Study	CELESTIAL TRIAL
	• Women of childbearing potential (premenopausal women capable of becoming pregnant and women who were amenorrheic for ≥12 months possibly due to prior chemotherapy, anti-oestrogens, ovarian suppression, low body weight or other reasons) must not have been pregnant at screening
Settings and location where the data were collected	The CELESTIAL trial was conducted in the secondary care setting in 19 countries: • Europe: Belgium, France, Germany, Ireland, Italy, Poland, Romania, Spain, The Netherlands, Turkey and United Kingdom • North America (United States of America [USA] and Canada) • Australia and New Zealand • Asia: Hong Kong, Republic of Korea, Singapore and Taiwan
Trial drugs	• Experimental Arm: Cabozantinib 60 mg tablet once daily • Comparator Arm: Matched placebo In addition, best supportive care was provided, based on the following general guidelines: • Analgesia and the management of AEs due to analgesia • Treatment of liver decompensation in patients with non-neoplastic liver disease • Antibiotics to treat infection, such as peritonitis and pneumonia • Provision of nutritional support and psychological support, including the management of depression and anxiety with medication and/or counselling • Transfusions to maintain haemoglobin levels, as clinically indicated (but not the use of erythroid growth factors).
Permitted and disallowed concomitant medication	The use of any of the following medications was permitted if required, during the trial: • Antiemetics and anti-diarrhoeal medications • Granulocyte colony-stimulating factors (except for prophylactic use before initial treatment with study drug) • Hormone replacement and short-term systemic steroid treatment • Low-doses of aspirin for cardio protection (per local guidelines), of warfarin (≤1 mg/day) and of low molecular-weight heparin • Antiviral therapy for active HBV infection. The use of any the following was not permitted in patients receiving study drug: • Any investigational agent or medical device • Any drug or herbal product specifically for the treatment of HCC • Therapeutic doses of oral anticoagulants (e.g., Warfarin [>1 mg/day] or warfarin-related agents, thrombin or factor Xa inhibitors) or antiplatelet agents (e.g., Clopidogrel); interferon • Liver-directed local anticancer therapy or systemic anti-tumour therapies • Erythropoietic-stimulating agents (e.g., Epoetin alfa and darbepoetin alfa)

Study	CELESTIAL TRIAL
	Palliative external radiation to bone metastasis or skin/subcutaneous metastasis was permitted during the trial but was discouraged unless medically unavoidable.
Primary outcome	• Overall Survival (OS) [Time Frame: Up to 45 months]
Secondary outcomes	• Progression-Free Survival (PFS) [Time Frame: Up to 45 months] • Objective Response Rate (ORR) [Time Frame: ORR is measured by radiologic assessment every 8 weeks after randomisation until disease progression or discontinuation of study treatment (up to 45 months)]
Exploratory endpoints	• HRQoL using EQ-5DL questionnaire • Safety and tolerability: evaluation of AEs, serious AEs (SAEs), deaths, clinical laboratory tests (haematology, serum chemistry and urinalysis), physical examination, vital signs, ECOG PS, 12-lead electrocardiogram (ECG) and the TTD in months (date of decision to discontinue study drug – date of first dose +1)/30.4375

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; ECOG, Eastern Cooperative Oncology Group; HbA1c, haemoglobin A1c; HBV, hepatitis B infection; HCC, hepatocellular carcinoma; ORR, objective Response Rate; OS, overall Survival; PFS, progression-free survival; PS, performance status; ULN, upper limit of normal; UPCR, urine protein/creatinine ratio; USA, United States of America.
Source: Exelixis, 2018 (44).

B.2.3.2 Baseline characteristics and demographics

The intention to treat (ITT) population included all patients randomised to receive study drug prior to the cut-off date for the second interim analysis, i.e., 1 June 2017, regardless of whether they received any/the correct study drug (6). The ITT population comprised 470 patients in the cabozantinib group and 237 patients in the placebo group. Demographic and baseline characteristics in the ITT population were well balanced between the treatment groups. Overall, almost half of the study population were ≥65 years of age (49%) and 82% were male. Most patients were White (56%) or Asian (34%). ECOG performance status (PS) was 0 in 53% of patients and 1 in 47% of patients; a single patient in the cabozantinib group had an ECOG PS of 1 at screening and 2 at baseline (44).

Stratification factors in the ITT population were also balanced between the treatment groups (Table 9). The stratification factors consisted of the following:

- Etiology of disease (hepatitis B virus [HBV] (HBV [with or without hepatitis C virus (HCV)], HCV [without HBV], or Other)
- Geographic region (Asia, Other Regions)
- Presence of extrahepatic spread of disease and/or macrovascular invasion (Yes, No)

The majority of patients were enrolled in Europe or North America (72%), 25% were enrolled in Asia and 4% in Australia/New Zealand. HBV [with or without HCV] was present in 38% of patients, 21% had HCV (without HBV) and 40% had HCC of another aetiology. Most patients (78%) had extrahepatic disease spread and/or macrovascular invasion (44).

Table 9: Baseline characteristics of patients in the CELESTIAL trial

Study	CELESTIAL Trial	
Baseline patient and disease characteristics	**Cabozantinib (n=470)**	**Placebo (n=237)**
Age, years, Median (range)	64 (22, 86)	64 (24, 86)
Sex, n (%) **Male**	379 (81)	202 (85)

Study	CELESTIAL Trial	
Baseline patient and disease characteristics	**Cabozantinib (n=470)**	**Placebo (n=237)**
Race, n (%)		
White	264 (56)	130 (55)
Asian	159 (34)	82 (35)
Black or African American	8 (2)	11 (5)
Other	8 (2)	2 (1)
Not reported	31 (7)	12 (5)
Geographic region, n (%)		
Europe	231 (49)	108 (46)
Asia	116 (25)	59 (25)
North America (USA/Canada), n (%)	108 (23)	59 (25)
Australia/New Zealand	15 (3)	11 (5)
ECOG PS, n (%)		
0 (normal activity, asymptomatic)	245 (52)	131 (55)
1 (fully ambulatory, symptomatic)	224 (48)	106 (45)
2 (in bed <50% of time, ambulatory and capable of self-care but not work activities)	1 (<1)	0
Aetiology at baseline, according to the CRF, n (%)		
Dual HBV and HCV	8 (2)	4 (2)
HBV	178 (38)	89 (38)
HCV	113 (24)	55 (23)
Alcohol related	112 (24)	39 (16)
NASH	43 (9)	23 (10)
Other/unknown	99 (21)	63 (27)
Child-Pugh A status, according to the CRF, n (%)		
A (score 5–6)	462 (98)	235 (99)
B (score 7–9)	7 (1)	2 (0.8)
Missing	1 (0.2)	0
Baseline disease, according to the CRF, n (%)		
Extrahepatic spread	369 (79)	182 (77)
Macrovascular invasion	129 (27)	81 (34)
AFP ≥400 ng/mL, n (%)	192 (41)	101 (43)
Prior systemic non-radiation anticancer regimens for advanced HCC, n (%)		
0	3 (0.6)	0
1	335 (71)	174 (73)
2	130 (28)	62 (26)
≥3	2 (0.4)	1 (0.4)
Median (range)	1 (0, 3)	1 (1, 3)
Duration of prior sorafenib for HCC, months, median (range)	5.32 (0.3, 70.0)	4.80 (0.2, 76.8)
<1month, n (%)	11 (2)	8 (3)
≥1 to <3 months, n (%)	117 (25)	54 (23)
≥3 to <6 months, n (%)	130 (28)	67 (28)
≥6 months, n (%)	211 (45)	108 (46)
Time from progression on sorafenib as most recent prior systemic agent, months, median (range)	n=322 1.61 (0.1, 28.3)	n=166 1.66 (0.2, 69.4)
Prior local liver-directed therapy (including transarterial chemoembolisation [TACE]), for HCC, n (%)	209 (44)	113 (48)
Prior TACE, for HCC, n (%)	203 (43)	111 (47)

Baseline was considered the last observation prior to randomisation; multiple aetiologies could be reported for each patient.

Abbreviations: AFP, alpha-fetoprotein; CRF, case report form; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; ITT, intention to treat; NASH, non-alcoholic steatohepatitis; PD-1, programmed cell death-1; PD-L1, programmed cell death ligand 1.
Source: Abou-Alfa et al., 2018 (6), Exelixis, 2018 (44).

CELESTIAL trial population compared with a typical UK population with HCC

The overall study population in the CELESTIAL trial were largely similar to a typical population of patients with advanced HCC in the UK, based on a retrospective national audit including data from 448 patients from 15 hospitals who received first-line systemic therapy with sorafenib for HCC (Table 10) (45).

Due to the inclusion criteria in the CELESTIAL trial, a higher proportion of patients participating in this study had an ECOG PS of 0 and more patients had Child-Pugh status A compared with a typical population of patients with HCC (Table 10).

	Observational data (N=448)	CELESTIAL trial (overall)* (N=707)
Age, years **Median (range)**	68 (17.0, 89.0)	64.0 (22, 86)
Sex, n (%)		
Male	325 (72.5)	581 (82.2)
Missing	57 (12.7)	0
ECOG PS, n (%) **0** **1** **2** **3** **Missing**	 117 (26.1) 218 (48.7) 94 (21.0) 6 (1.3) 13 (2.9)	 376 (53.2)) 330 (46.7) 1 (0.1) † 0 0
Disease characteristics		
Child-Pugh status, n (%) **A** **B** **C** **Missing**	 343 (76.6) 72 (16.1) 2 (0.4) 31 (6.9)	 697 (98.6) 9 (1.3) 0 1 (0.1)
Presence of extrahepatic spread, n (%) **Yes** **Missing**	 172 (38.4) 7 (1.6)	 551 (77.9) –

	Observational data (N=448)	CELESTIAL trial (overall)* (N=707)
Presence of micro/vascular invasion, n (%)		
Yes	91 (20.3)	210 (29.7)
Missing	196 (43.8)	–
AFP ≥400 ng/mL, n (%)**	141 (31.5)	293 (41.4)
Missing	80 (17.9)	0
Aetiology of disease, n (%)		
HBV	55 (12.3)	267 (37.8)
HCV	70 (15.6)	168 (23.8)
Alcohol related	110 (24.6)	151 (21.4)
Previous local therapy, n (%)	190 (42.4)	324 (45.8)

A higher proportion of patients in the CELESTIAL trial had extensive metastatic disease at baseline, with almost double the proportion of patients with extrahepatic spread. In addition, a higher proportion of patients participating in the CELESTIAL trial had HBV and/or HCV (Table 10).

Table 10: Baseline and disease characteristics of a typical population of patients with advanced HCC in the UK (based on observational data) and participants in the CELESTIAL trial

	Observational data (N=448)	CELESTIAL trial (overall)* (N=707)
Age, years		
Median (range)	68 (17.0, 89.0)	64.0 (22, 86)
Sex, n (%)		
Male	325 (72.5)	581 (82.2)
Missing	57 (12.7)	0
ECOG PS, n (%)		
0	117 (26.1)	376 (53.2))
1	218 (48.7)	330 (46.7)
2	94 (21.0)	1 (0.1) †
3	6 (1.3)	0
Missing	13 (2.9)	0
Disease characteristics		
Child-Pugh status, n (%)		
A	343 (76.6)	697 (98.6)
B	72 (16.1)	9 (1.3)
C	2 (0.4)	0
Missing	31 (6.9)	1 (0.1)
Presence of extrahepatic spread, n (%)		
Yes	172 (38.4)	551 (77.9)
Missing	7 (1.6)	–

	Observational data (N=448)	CELESTIAL trial (overall)* (N=707)
Presence of micro/vascular invasion, n (%)		
Yes	91 (20.3)	210 (29.7)
Missing	196 (43.8)	–
AFP ≥400 ng/mL, n (%)**	141 (31.5)	293 (41.4)
Missing	80 (17.9)	0
Aetiology of disease, n (%)		
HBV	55 (12.3)	267 (37.8)
HCV	70 (15.6)	168 (23.8)
Alcohol related	110 (24.6)	151 (21.4)
Previous local therapy, n (%)	190 (42.4)	324 (45.8)

*Intention to treat population, according to the case report form (CRF) in the CELESTIAL trial
†A patient in the cabozantinib group had an ECOG PS of 1 at screening and 2 at baseline
** AFP ≥400 ng/mL defines a poorer prognostic group
Abbreviations: AFP, alpha-fetoprotein; ECOG, Eastern Cooperative Oncology Group; HBV, hepatitis B virus; HCV, hepatitis C virus; N/A, not available; PS, performance status.
Source: Abou-Alfa et al., 2018 (6), Exelixis, 2018 (44), King et al., 2017 (45)

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

B.2.4.1 Analysis populations

All efficacy analyses were conducted using data from the ITT population (44). The results from the second planned interim analysis are presented in this document. For the second interim analysis, the ITT population comprised all patients randomised to receive study drug as of the cut-off date for the second interim analysis, i.e., 1 June 2017, regardless of whether they received any/the correct study drug. (Table 11)

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 11: Analysis sets in the CELESTIAL trial

Analysis sets	Number of patients		
	Cabozantinib	Placebo	Total
ITT			
Overall population	**470**	**237**	**707**
Safety			
Overall population	**467***	**237**	**704**

* Three patients did not receive study drug
Abbreviations: ITT, intention to treat.
Source: Abou-Alfa et al., 2018 (6).

As of 18 September 2017, 773 patients had been enrolled in the trial (target sample size 760) and enrolment was closed (6).

B.2.4.2 Statistical analysis

An overview of the primary statistical analyses in the CELESTIAL trial is provided in Table 12 (44).

Sensitivity analyses

In addition to the primary analysis of PFS (PFS1), sensitivity analyses were undertaken (PFS2 and PFS3) that included defining additional clinical outcomes as events and evaluated the impact of informative censoring, an overview of which is shown in Table 13.

Table 12: Summary of the statistical analyses undertaken in the CELESTIAL trial

Hypothesis objective	Statistical analysis	Sample size, power calculation
The null hypothesis was that there was no difference in the duration of OS between the treatment groups (cabozantinib plus BSC versus placebo plus BSC) The alternative hypothesis was that there was a difference in the duration of OS between the treatment groups (cabozantinib plus BSC versus placebo plus BSC)	**Primary efficacy analyses** *Primary efficacy endpoint: OS* Analyses: Up to three analyses were planned: two interim analyses and a final analysis when approximately 50%, 75% and 100% of the total required number of deaths, respectively, were observed, i.e., 311, 466 and 621 deaths, respectively. Hypothesis testing was performed using the stratified log-rank test with a two-sided α=0.05. The stratification factors were the same as those used to stratify randomisation (IxRS data were used). Median duration of OS and the associated 95% confidence interval (CI) for each treatment group was estimated using the Kaplan-Meier method. The stratified HR and its 95% CI were estimated using a Cox proportional hazard model with treatment group as the independent variable and stratified by the randomisation/log-rank test stratification factors. Inflation of Type I error associated with interim analyses was controlled using a Lan-DeMets O'Brien-Fleming alpha-spending function. The calculated critical p-values (and observed hazard ratios [HR]) for rejecting the null hypothesis were 0.0031 (HR ≤0.70), 0.0183 (HR ≤0.80) and 0.044 (HR ≤0.84) for 311, 466 and 621 deaths (50%, 75% and 100% of deaths), respectively. The actual critical values were based on the actual number of events observed at the time of each analysis. The actual critical value for the first interim analysis was 0.0037 (321 deaths, 52% of the total required number of deaths) and for the second interim analysis was 0.021 (484 deaths, 78% of the total required number of deaths). 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 OS 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 OS was rejected in favour of cabozantinib. Formal futility analyses were not planned. *Secondary efficacy endpoint* PFS: investigator-determined radiographical progression according to RECIST 1.1 (only adequate tumour assessments [ATAs] were considered) or death.	The sample size was based on the primary efficacy endpoint (OS). A sample size of 760 patients and 621 events would provide 90% power for a two-sided log-rank test at 5% significance to detect a 31.6% increase in OS with cabozantinib compared with placebo (HR 0.76). Assuming a median OS of 8.2 months in the placebo group (based upon the placebo-controlled brivanib BRISK trial in patients who were previously treated with sorafenib (47)) and exponential distribution, this would correspond to median OS of 10.8 months in the cabozantinib group. The minimum observed effect that would result in statistical significance for OS at the two interim analyses and the final analysis were 42.1% improvement (HR 0.70, i.e. from 8.2 to 11.7 months), 25.7% improvement (HR 0.80, i.e. from 8.2 to 10.3 months) and 18.4% improvement (HR 0.84, i.e. from 8.2 to 9.7 months), respectively.

Hypothesis objective	Statistical analysis	Sample size, power calculation
	ORR: the proportion of patients with a CR or PR as the investigator-determined BOR in terms of tumour assessment category (CR, PR, stable disease, progressive disease or not evaluable) according to RECIST 1.1 that occurred prior to any censoring relevant for the primary analysis of PFS (see Table 13 for censoring details). Analysis of the secondary endpoints only took place if the result of either an interim analysis or the final analysis of OS achieved statistical significance compared with placebo. The hypotheses for PFS and ORR were tested in parallel; PFS was tested with a two-sided α=0.04 and ORR with a two-sided α=0.01. The primary analysis of PFS was performed in a similar manner to the primary analysis of OS For BOR, confirmation of response was required ≥28 days after the response was first observed. Hypothesis testing for ORR was performed using Fisher exact test. Analysis using the Cochran-Mantel-Haenszel (CMH) method to adjust for randomisation stratification factors was also performed. If the ORR was >10%, the duration of the objective response and time to the objective response were calculated. The duration of objective response (the time from the first documentation of objective response by the investigator, confirmed ≥28 days later, to disease progression or death due to any cause) was calculated using the Kaplan-Meier method with the dates of progression and censoring determined as described for the analysis of PFS. The time to objective response was the time from randomisation to the first documentation of objective response by the investigator, which was confirmed ≥28 days later. *Multiplicity* The multiplicity issue resulting from analysis of one primary endpoint, two secondary efficacy endpoints (PFS and ORR) and planning two interim analyses for testing OS was addressed by employing a fixed-sequence testing procedure, applying a modified Bonferroni procedure (dividing the α between the secondary endpoints), and implementing an α-spending function. *Exploratory endpoints* Safety was analysed descriptively. In general, other than for partial dates, missing data were not imputed. For patient reported outcomes. the change in EuroQol five-dimension five-level (EQ-5D-5L) questionnaire scores from first assessment to the end of the study were summarised descriptively at each post-baseline time point (every 4 weeks until week 25 and then every 8 weeks) and compared using a repeated-measures mixed-effects analysis. For the EQ-5D-5L	

Hypothesis objective	Statistical analysis	Sample size, power calculation
	index scores and EQ-5D-5L visual analogue scores (VAS), the mean change from baseline score to each post-baseline visit were summarised descriptively. A minimal important difference (MID) for these questionnaires in cancer patients were previously established as 0.06 - 0.08 for EQ-5D Index, and 7 for EQ-VAS (46)	

Abbreviations: BOR, best overall response; BSC, best supportive care; CR, complete response; EQ-5D-5L, EuroQol five-dimension five-level; EQ-VAS, EuroQol visual analogue scale MID, minimal important difference; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PR, partial response.
Source: Exelixis, 2018 (44).

Table 13: Event and censoring rules for the primary analysis of PFS (PFS1) and the sensitivity analyses (PFS2 and PFS3)

Analysis	PFS1		PFS2		PFS3	
Purpose	Primary		Sensitivity		Sensitivity	
Situation	Outcome	Date	Outcome	Date	Outcome	Date
No post-baseline assessment	Censored	Date of randomisation	Censored	Date of randomisation	Censored	Date of randomisation
Radiographical PD	Event	Date of PD	Event	Date of PD	Event	Date of PD
Death	Event	Date of PD	Event	Date of PD	Event	Date of PD
Subsequent systemic or local liver-directed NPACT	Censored	Date of last ATA on or prior to date of NPACT	Event	Date of NPACT	Censored	Date of last ATA on or prior to date of NPACT
Radiation (other than to bone)	Censored	Date of last ATA on or prior to date of radiation	Event	Date of radiation	Censored	Date of last ATA on or prior to date of radiation
Surgery to resect tumour lesions	Censored	Date of last ATA on or prior to date of surgery	Event	Date of surgery	Censored	Date of last ATA on or prior to date of surgery
Event after >2 missed ATAs (>126 days)	Censored	Date of last ATA prior to the missing visits	Censored	Date of last ATA prior to the missing visits	Censored	Date of last ATA prior to the missing visits
Treatment discontinuation due to clinical deterioration	NA	NA	Event	Date of determination	Event	Date of determination
No event by last ATA	Censored	Date of last ATA	Censored	Date of last ATA	Censored	Date of last ATA

Abbreviations: ATA, adequate tumour assessments; NPACT, non-protocol anticancer therapy; PD, progressive diseases; PFS, progression-free survival.
Source: Exelixis, 2018 (44).

B.2.4.3 Participant flow in the CELESTIAL trial

See the CONSORT diagram for the CELESTIAL trial in Appendix D (6).

B.2.5. Quality assessment of the relevant clinical effectiveness evidence

A quality assessment of the CELESTIAL trial is summarised in Table 14. The CELESTIAL 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 14: Quality assessment results for the CELESTIAL trial

Trial	The CELESTIAL 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.

B.2.6. Clinical effectiveness results of the relevant trials

B.2.6.1 Primary endpoint: OS

The data presented are from the second interim analysis, planned for when 75% of the total number of required deaths to adequately power the trial (621 deaths), i.e., 466 deaths, had occurred (6, 44). At the cut-off date for the second interim analysis (1 June 2017), 484 deaths in the overall population had been reported, representing 78% of the total number of deaths required. The median duration of follow-up for OS was 22.9 months. Cabozantinib significantly reduced the risk of death by 24% compared with placebo (HR 0.76 [95% CI: 0.63, 0.92]; stratified log-rank p-value 0.005) increasing the median OS by 2.2 months (10.2 versus 8.0 months) (Table 15; Figure

4). The landmark estimate of the proportion of patients alive at 12 months was 46% in the cabozantinib group compared with 34% in the placebo group (6).

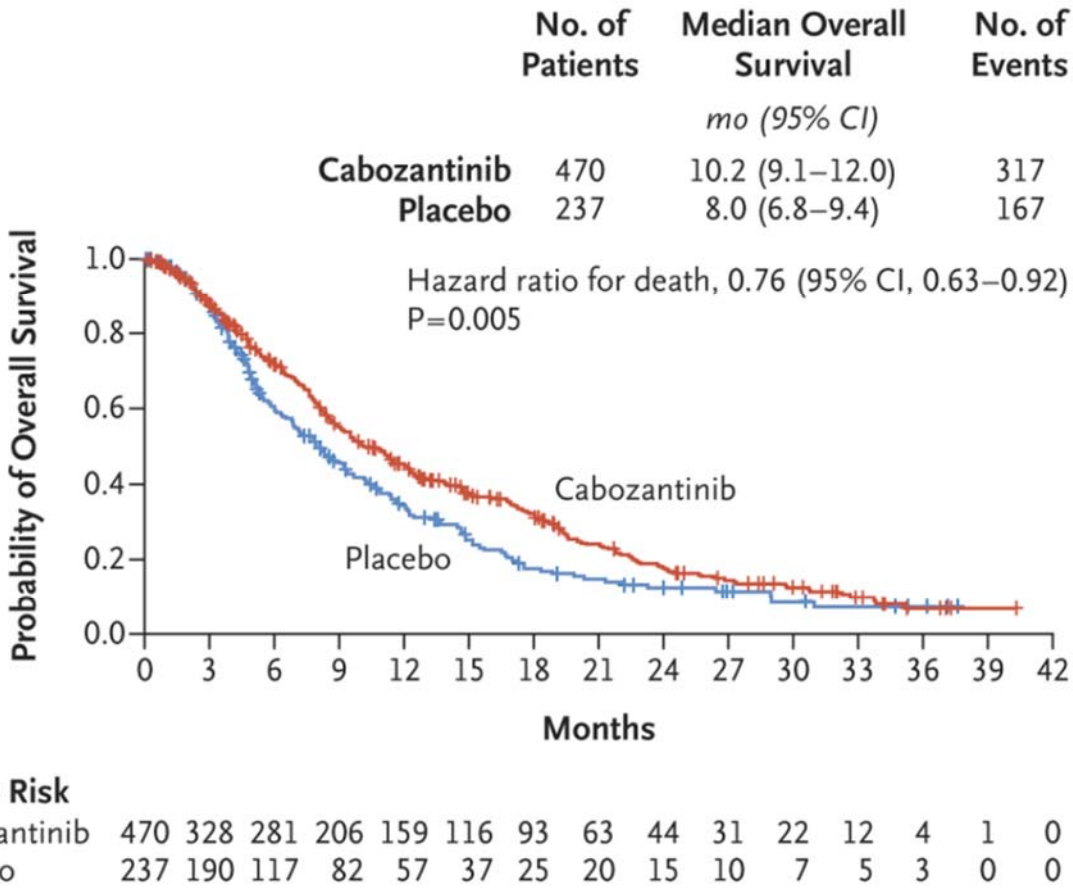
Thus, the null hypothesis that there was no difference in the duration of OS between the treatment groups (cabozantinib plus BSC versus placebo plus BSC) was rejected as a result of the second interim analysis. As a result of this no further analyses of OS were planned.

Table 15: The CELESTIAL trial: duration of OS (ITT; second planned interim analysis)

	Cabozantinib (n=470)	Placebo (n=237)
Patients, n (%)		
Censored	153 (33)	70 (30)
Death	317 (67)	167 (70)
Duration of OS (months)		
Median (95% CI)	10.2 (9.1, 12.0)	8.0 (6.8, 9.4)
Range	0.1, 40.3+	0.03+, 37.6+
Critical p-value to reject null hypothesis of equal OS	0.02	
Observed p-value (stratified log-rank test)	0.005	
HR (95% CI; stratified)	0.76 (0.63, 0.92)	
Observed p-value (unstratified log-rank test)	0.0072	
HR (95% CI; unstratified)	0.77 (0.64, 0.93)	

+ indicates a censored observation
Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; OS, overall survival.
Source: Abou-Alfa et al., 2018 (6). Exelixis, 2018 (44).

Figure 4. The CELESTIAL trial: OS with cabozantinib versus placebo – Kaplan-Meier plot (ITT population; second planned interim analysis, adjusted)



Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; OS, overall survival.
Source: Abou-Alfa et al., 2018 (6).

B.2.6.2 Secondary endpoint: PFS

Analysis of PFS was conducted in the ITT population at the time of the primary analysis of the primary endpoint OS, i.e., at the time of the second planned interim analysis, due to the significant result for the primary endpoint (6, 44). In the pre-specified primary analysis of the secondary efficacy endpoint PFS, PFS was defined as the time from randomisation to investigator-determined radiographical progression according to RECIST 1.1 or death due to any cause in the ITT population.

Cabozantinib significantly reduced the risk of disease progression/death by 56% compared with placebo (HR 0.44 [95% CI: 0.36, 0.52]; stratified log-rank p-value <0.0001) increasing median PFS by 3.3 months (5.2 versus 1.9 months) at the time of the second planned interim analysis (Table 16; Figure 5). The landmark estimate of

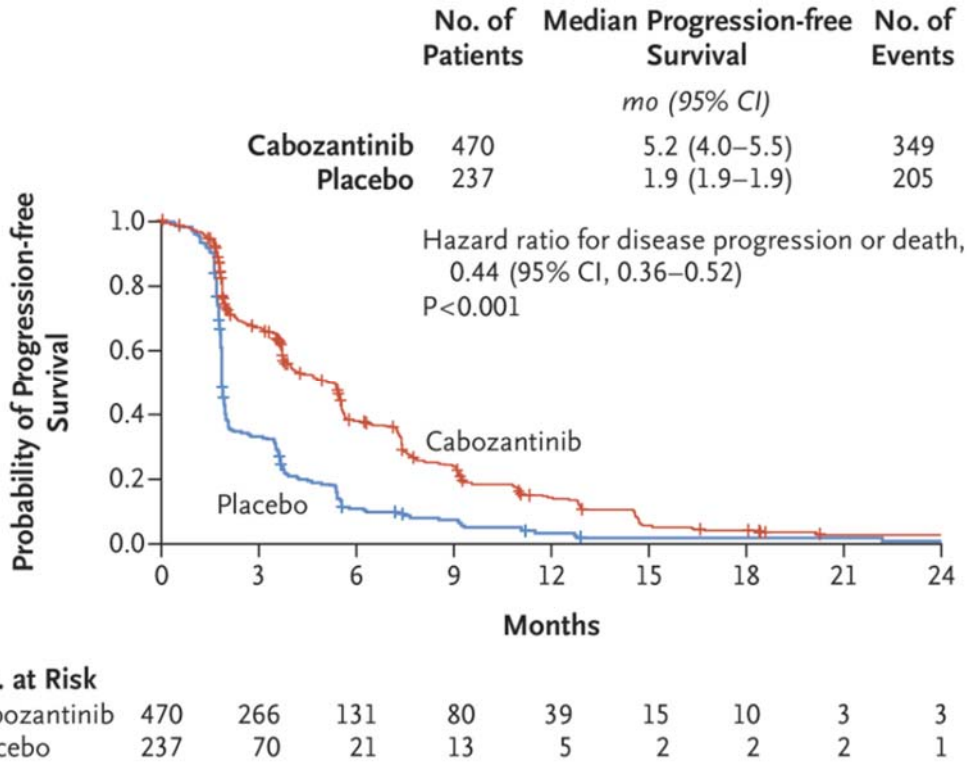
the proportion of patients alive and progression-free at 12 months was 15% in the cabozantinib group compared with 3% in the placebo group.

Table 16: The CELESTIAL trial: PFS (investigator assessed; ITT population; second interim analysis)

	Cabozantinib (n=470)	Placebo (n=237)
Number (%) of patients		
Censored	121 (26)	32 (14)
Event	349 (74)	205 (86)
Death	65 (14)	19 (8.0)
PD	284 (60)	186 (78)
Duration of PFS (months)		
Median (95% CI)	5.2 (4.0, 5.5)	1.9 (1.9, 1.9)
Range	0.03+, 33.2	0.03+, 25.5+
Critical p-value to reject null hypothesis of equal PFS	0.04	
Observed p-value (stratified log-rank test)	<0.0001	
HR (95% CI; stratified)	0.44 (0.36, 0.52)	
Observed p-value (unstratified log-rank test)	<0.0001	
HR (95% CI; unstratified)	0.46 (0.38, 0.55)	

+ indicates a censored observation
Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; PD, progressive disease; PFS, progression-free survival**.**
Source: Abou-Alfa et al., 2018 (6). Exelixis, 2018 (44).

Figure 5. The CELESTIAL trial: PFS – Kaplan-Meier plot (investigator assessed; ITT population; second interim analysis, adjusted)



Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; PFS, progression-free survival.
Source: Abou-Alfa et al., 2018 (6).

The robustness of the significant improvement in PFS with cabozantinib compared with placebo was confirmed in the unadjusted analysis and in sensitivity analyses. The results of two sensitivity analyses (PFS2 and PFS3; data not used in the economic model) in which PFS was defined using additional clinical outcomes as events and which also evaluated the impact of informative censoring were similar to those in the primary analysis (Table 17).

Table 17: The CELESTIAL trial: results of sensitivity analyses for PFS (investigator assessed; ITT population; second interim analysis)

PFS analysis	Cabozantinib (n=470)		Placebo (n=237)		Cabozantinib versus placebo	
	Events, % (n)	Mean, months	Events, % (n)	Mean, months	HR (95% CI) stratified	p-value log-rank test, stratified
Primary analysis	74 (349)	5.2	86 (205)	1.9	0.44 (0.36, 0.52)	<0.0001
Sensitivity analyses						
PFS2	80 (374)	4.4	89 (211)	1.9	0.46 (0.38, 0.55)	<0.0001

PFS analysis	Cabozantinib (n=470)		Placebo (n=237)		Cabozantinib versus placebo	
	Events, % (n)	Mean, months	Events, % (n)	Mean, months	HR (95% CI) stratified	p-value log-rank test, stratified
PFS3	76 (356)	4.7	87 (207)	1.9	0.44 (0.37, 0.53)	<0.0001

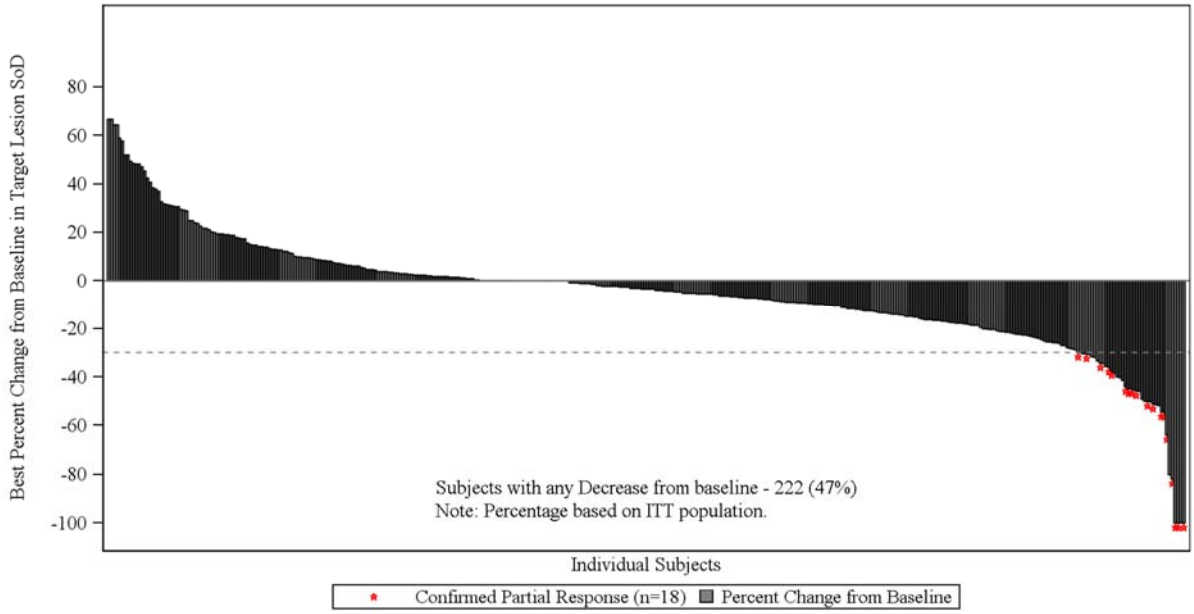
Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intention to treat; OS, overall survival; PFS, progression-free survival.
Source: Abou-Alfa et al., 2018 (6) Exelixis, 2018 (44)

B.2.6.3 Secondary endpoint: ORR

Analysis of the secondary efficacy endpoint ORR (investigator-determined CR or PR according to RECIST 1.1) was conducted in the ITT population at the time of the primary analysis of OS, i.e., at the time of the second planned interim analysis, due to the significant result for the OS (6, 44).

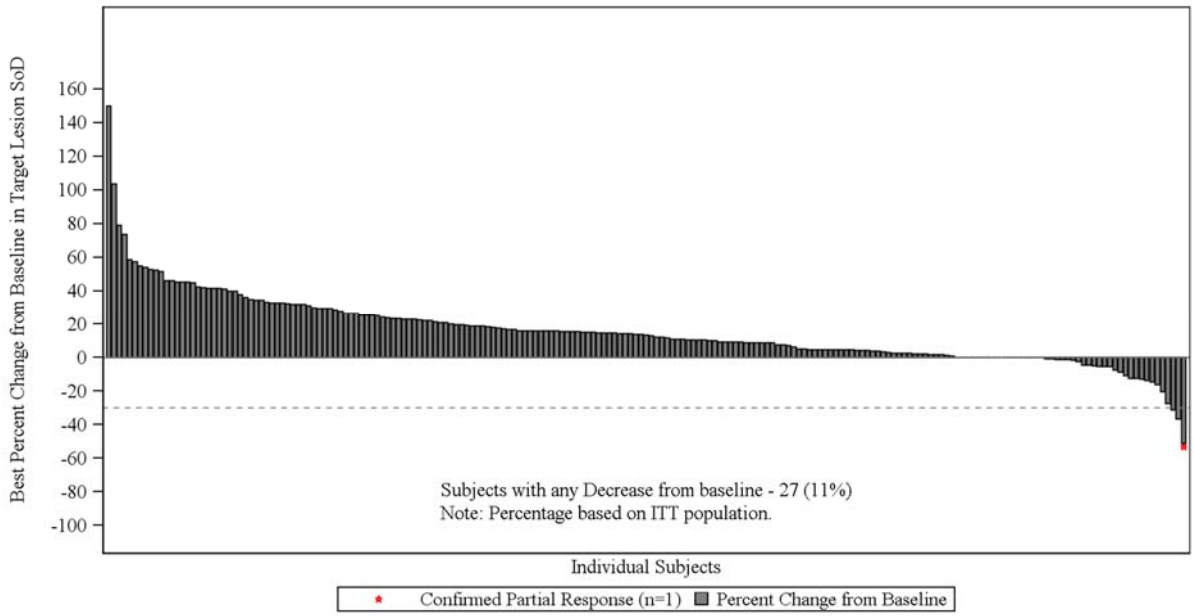
The best percentage change from baseline in tumour target lesion size (investigator-determined according to RECIST 1.1) is depicted in Figure 6 (cabozantinib) and Figure 7 (placebo). Post-baseline reduction in the sum of target lesion diameters (SoD) was observed in 47% of subjects in the cabozantinib arm and 11% in the placebo arm. The waterfall plots do not include subjects which lack of evaluable post-baseline assessment, censoring (per PFS rules) before first evaluable post-baseline assessment, lack of target lesions, and/or incomplete or unevaluable target lesion assessment. Data from time points after the first date of any of the censoring events defined for the primary PFS analysis were also excluded from the plots.

Figure 6. Waterfall plot of best percentage change in tumour target lesion size from baseline per Investigator; Cabozantinib arm (ITT population, subjects with a baseline and post-baseline target lesion assessment, N = 388)



Abbreviations: ITT, intention to treat; SoD, sum of target lesion diameters.
Source: Exelixis, 2018 (44).

Figure 7. Waterfall plot of best percentage change in tumour target lesion size from baseline per investigator; Placebo arm (ITT population, subjects with a baseline and post-baseline target lesion assessment, N = 205)



Abbreviations: ITT, intention to treat; SoD, sum of target lesion diameters.
Source: Exelixis, 2018 (44).

The results for BOR clearly demonstrate a higher disease control rate with cabozantinib compared with placebo (64% versus 33%). Cabozantinib was associated

with a significantly higher ORR than placebo (odds ratio [OR] 9.4 [95% CI 1.2, 71.0]; stratified Cochran-Mantel-Haenszel (CMH) p-value 0.0086). As no patient in either treatment group had a CR, these results reflect the significantly higher PR rate with cabozantinib compared with placebo (4% versus 0.4%). As would be expected due to the significantly higher ORR with cabozantinib, cabozantinib was also associated with a lower rate of progressive disease (PD) compared with placebo (21% versus 55%) (Table 18).

Table 18: The CELESTIAL trial: ORR for cabozantinib versus placebo (investigator-determined; ITT population; second interim analysis)

	Cabozantinib n=470	Placebo n=237
BOR, n (%)		
Confirmed CR	0	0
Confirmed PR	18 (4)	1 (0.4)
SD	282 (60)	78 (33)
Unconfirmed CR	0	0
Unconfirmed PR	13 (3)	2 (0.8)
PD	98 (21)	131 (55)
Unable to evaluate/missing	72 (15)	27 (11)
No baseline assessment	0	0
No post-baseline assessments	65 (14)	22 (9)
No qualifying post-baseline assessment on or before primary PFS analysis censoring or event	7 (1)	5 (2)
ORR [CR + PR], n (%)	18 (4)	1 (0.4)
95% CI	(2.3, 6.0)	(0.0, 2.3)
Treatment difference (cabozantinib – placebo) (95% CI)	3.4 (1.49, 5.33)	
Critical p-value to reject null hypothesis of equal ORR	0.01	
Observed p-value (stratified CMH test)	0.0086	
Odds ratio, stratified (95% CI)	9.4 (1.2, 71.0)	
Observed p-value (unstratified Fishers exact test)	0.0059	
Odds ratio, unstratified (95% CI)	9.4 (1.2, 70.8)	

Abbreviations: BOR, best overall response; CI, Confidence Interval; CMH, Cochran-Mantel-Haenszel; CR, complete response; ITT, intention to treat; ORR, objective response rate; OS, overall survival; PD, progressive disease; PFS, progression-free survival; PR, partial response; SD, stable disease
Source: Abou-Alfa et al., 2018 (6), Exelixis, 2018 (44)

B.2.6.4 Exploratory endpoints

Time to treatment discontinuation

In the CELESTIAL trial, patients received cabozantinib for almost twice as long as patients received placebo: the median duration of exposure at the time of the planned second interim analysis of OS (cut-off date 1 June 2017) was 3.8 months (range 0.1, 37.3) in the cabozantinib group compared with 2.0 months (range 0.0, 27.2) in the placebo group (6, 44).

Safety

Data regarding AEs are reported in Section B.2.10.

Patient reported outcomes

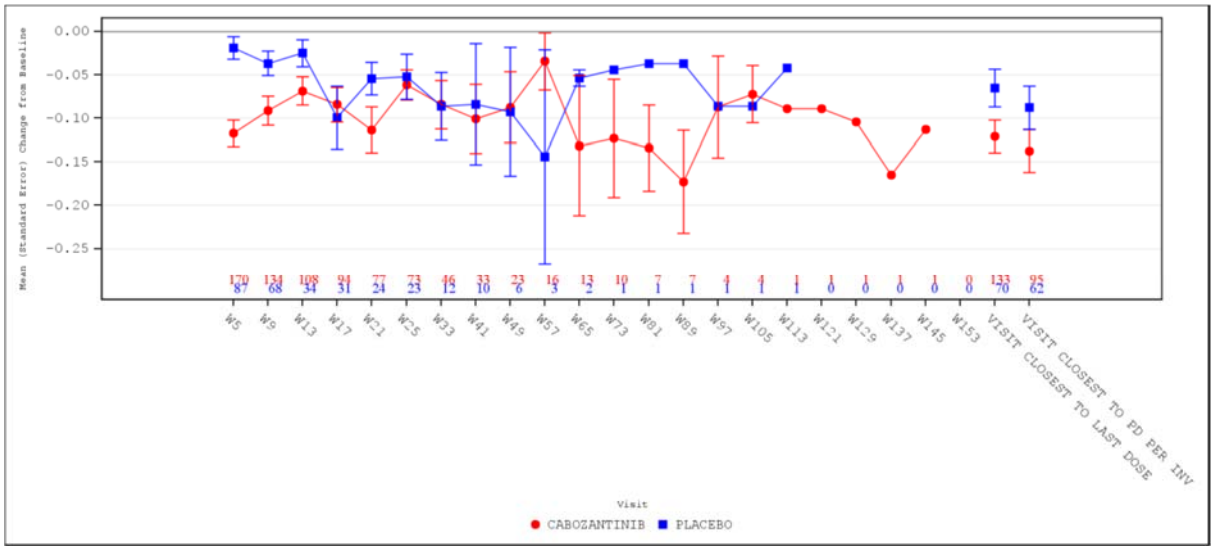
To assess symptom burden and patients' HRQoL EQ-5D-5L were collected in the CELESTIAL study. Questionnaires were completed by patients at baseline, and post-baseline assessments were collected every 4 weeks until week 25 and then every 8 weeks, on the same schedule as tumour CT/MRI assessments. Assessments continued regardless of whether study treatment was given, reduced, interrupted, or discontinued until the later of 8 weeks after radiographic progression per Investigator or the date of the decision to discontinue study treatment. Subjects were not to receive medical results prior to completing the questionnaire (44).

Completion rates (number of subjects who completed all questions/number of expected subjects still on study at each visit) remained above 85% in each treatment arm through Week 33. Beyond Week 33, there were fewer than 20 subjects in the placebo arm (44).

At baseline, mean EQ-5D-5L scores were higher (with corresponding lower utility scores) for cabozantinib compared to placebo across all five health domains (mobility, self-care, usual activity, pain/discomfort, anxiety/depression, and utility). At baseline, mean EQ-5D Index scores were 0.792 in the cabozantinib arm and 0.855 in the placebo arm. At baseline, mean EQ-VAS scores were 73.5 in the cabozantinib arm and 76.1 in the placebo arm.

The change from baseline for EQ-5D Index is shown in Figure 8. At week 5, there was a statistically significant reduction in mean health utility scores for cabozantinib compared with placebo (difference of −0.097). This decrement in the cabozantinib (versus placebo) group remained statistically significant, but below the MID, at each visit from week 5 to week 21. During weeks 25–81, the difference ceased to be statistically significant and switched to favouring cabozantinib at weeks 33, 49 and 65. The confidence intervals around the scores were wide, however, making the true clinical significance of the difference difficult to discern.

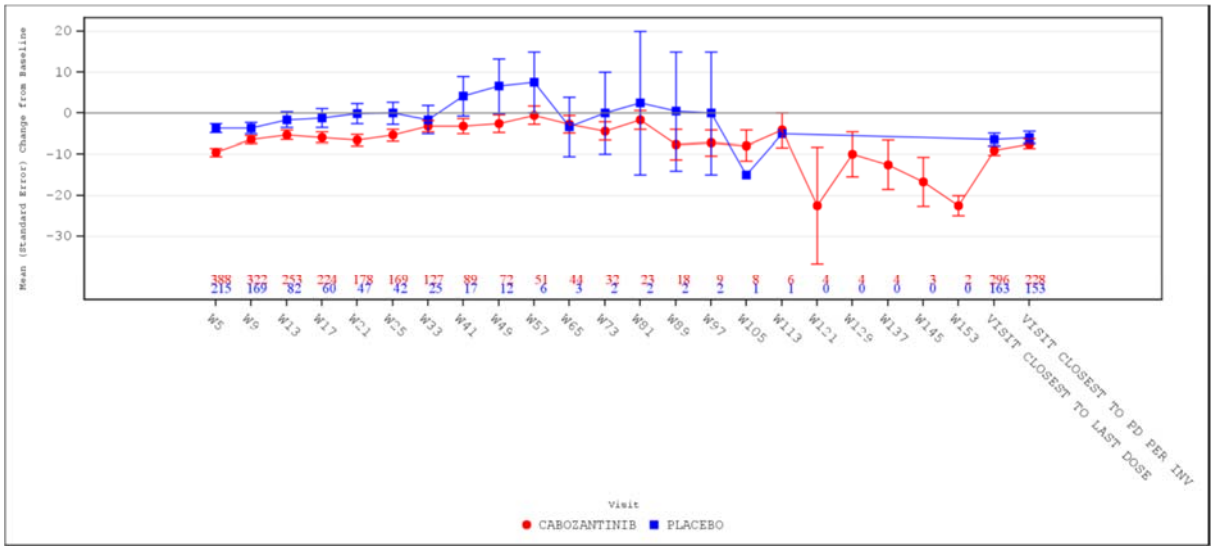
Figure 8: Mean change from baseline of EQ-5D Index score (Countries in which EQ-5D Index Is Validated)



Abbreviations: EQ-5D, EuroQol five-dimension

The change from baseline for EQ-VAS is shown in Figure 9. All treatment differences in mean change from baseline EQ-VAS values were <7 through Week 33. Beyond this time point, there were fewer than 20 subjects in the placebo arm.

Figure 9: Mean change from baseline of EQ-VAS score



Abbreviations: EQ-VAS, EuroQol visual analogue scale

Table 19 shows the repeated measures analysis. There was a potentially clinically meaningful treatment difference in favour of placebo for EQ-5D Index (effect size -0.319). Effect size differences ≥ 0.3 were regarded as likely to be clinically relevant (48, 49). There was no clinically meaningful treatment difference in effect size for EQ-VAS.

Table 19: EQ-VAS and EQ-5D Index Scores: Change from baseline, repeated-measures analysis

	EQ-5D Index	EQ-VAS
Cabozantinib n (N = 470)	178	398
Cabozantinib least square means (SE)	-0.11 (0.020)	-8.30 (1.100)
Placebo n (N = 237)	90	216
Placebo least square means (SE)	-0.05 (0.022)	-3.87 (1.319)
Difference in mean change	-0.057	-4.432
Pooled SD	0.179	17.826
P-value	<0.0001	0.0002
Effect size	-0.319	-0.249

Abbreviations: EQ-5D, EuroQol five-dimension; EQ-VAS, EuroQol visual analogue scale; SE, standard error.

B.2.7. Subgroup analysis

There are no subgroups of interest as the target population is the full marketing authorisation. In an ad hoc subgroup analysis, subjects whose only prior therapy for HCC was sorafenib also showed an OS benefit. Subgroup analyses demonstrated a generally consistent OS and PFS benefit for cabozantinib treated patients in all subgroups comprising at least 20 patients. There were too few responders to interpret ORR subgroup analyses. The CELESTIAL study was not powered to assess differential patient response to treatment in subgroups.

More detailed results of the subgroup analysis are provided in Appendix E.

B.2.8. Meta-analysis

No meta-analysis was carried out, as the only two trials identified as relevant to the decision problem were the CELESTIAL trial that compared cabozantinib with placebo, and the RESORCE trial that compared regorafenib with placebo.

B.2.9. Indirect and mixed treatment comparisons

In the absence of a head-to-head trial comparing cabozantinib with regorafenib, the following ITCs have been conducted to estimate the relative efficacy of cabozantinib versus regorafenib:

- One based on Bucher et al. (50), and
- The other being a matching-adjusted indirect comparison (MAIC).

The CELESTIAL and RESORCE trials were identified as the only relevant trials to perform the indirect comparisons and both trials shared a common comparator treatment, placebo. The summary of these trials is included in Table 20.

B.2.9.1 Identification of studies

The systematic literature review (SLR) described in Appendix D, was used to identify all potential studies that may have been relevant for indirect comparison with cabozantinib.

Table 20: Summary of the trials used to carry out the indirect treatment comparison

Trial reference	CELESTIAL	RESORCE
Intervention (N)	Cabozantinib (60 mg qd) plus BSC (470)	Regorafenib (160 mg qd) plus BSC (379) - once daily during weeks 1–3 of each 4-week cycle
Comparator (N)	Placebo plus BSC (237)	Placebo plus BSC (194) - once daily during weeks 1–3 of each 4-week cycle
Study initiation and completion (years)	26 September 2013 – 01 June 2017 (data cut-off date)	May 2013 – Feb 2016 (primary completion date)
Phase	III	III
Patient population (ITT)	Sorafenib tolerant and intolerant; second and third-line patients (CELESTIAL inclusion criteria listed in Table 8)	Sorafenib tolerant, second-line patients only
Method of blinding	Double-blind	Double-blind
Randomisation	2:1, stratified by etiology of disease (HBV [with or without HCV], HCV [without HBV], or Other), geographic region (Asia, Other Regions), and presence of extrahepatic spread of disease and/or macrovascular invasion (yes versus no).	2:1, stratified by geographical region (Asia versus rest of world), macrovascular invasion (yes versus no), extrahepatic disease (yes versus no), α-fetoprotein concentration (<400 ng/mL versus ≥400 ng/mL), and ECOG performance status (0 versus 1).
Study centres	Multicentre (Europe, North America, Australia, New Zealand, Asia)	Multicentre (Europe, North America, Australia, South America, Asia)
Median follow-up duration	22.9 months	7.0 months
Patients censored for OS (%)	32%	37%

Abbreviations: BSC, best supportive care; ECOG, Eastern Cooperative Oncology Group; ITT, intention to treat; N, number of participants; OS, overall survival; qd, once a day.
Source: Abou-Alfa et al., 2018 (6), Exelixis, 2018 (44), Finn et al, 2018 (51)

B.2.9.2 Indirect treatment comparison based on Bucher et al methodology

An ITC based on the approach used by Bucher et al. (50) was performed to estimate the relative efficacy of cabozantinib versus regorafenib, in accordance with the decision problem outlined in Table 1. The principal assumption of the Bucher ITC is that the relative efficacy of the treatments included in the comparison is the same in all trials included in the indirect comparison. To satisfy this assumption, the trials need to be comparable in terms of study design and patient characteristics. For this analysis, however, it should be noted that the Bucher approach is limited by the fact that the ITT population results for the overall population of CELESTIAL trial which

included second and third-line patients would be compared against the overall population of the RESORCE trial which includes second-line patients only.

Comparison of trial design and patient characteristics

Both trials (CELESTIAL and RESORCE) were phase III, multicentre, double-blind RCTs, conducted over similar durations and in similar geographical locations suggestive of consistent clinical practices across both trials. However, the trials populations differed in several baseline characteristics with differences in the ethnic mix, region, ECOG performance status, number of prior treatments and duration of prior sorafenib treatment between the trial populations. A comparison of baseline characteristics showed that, on average, patients enrolled in the CELESTIAL trial had a shorter duration of prior sorafenib treatment than patients in RESORCE (8 versus 12 months). Additionally, patients in CELESTIAL were less likely to have an ECOG PS of 0 (53% versus 66%), more likely to be white (56% versus 36%), and less likely to be in the Asia geographical region (25% versus 38%). The baseline characteristics from CELESTIAL and RESORCE trial are presented in Table 21.

Table 21. Comparison of baseline characteristics of subjects enrolled in CELESTIAL and RESORCE

	CELESTIAL	RESORCE
Treatment (N)	**Cabozantinib (N = 470)**	**Regorafenib (N = 374)**
Age under 65	51	55
Female	18	12
Asia geographical region	25	38
ECOG status 0	53	66
Child-Pugh class A	100	98
Mean duration of sorafenib treatment (months)	8	12
Extrahepatic disease	78	72
Macrovascular invasion	30	29
Hepatitis B aetiology	38	38
Alcohol use aetiology	22	25
Hepatitis C aetiology	24	21
AFP > 400ng/mL	41	43
White (%)	56	36

Abbreviations: AFP; alpha fetoprotein; ECOG, Eastern Cooperative Oncology Group; N, number of patients.
Source: Abou-Alfa et al., 2018 (6), Exelixis, 2018 (44), Finn et al, 2018 (51).

Methodology

The effect of cabozantinib relative to regorafenib was estimated using the method for adjusted indirect comparison developed by Bucher et al. (50). The method applies aggregate data from the CELESTIAL and RESORCE trials, with the placebo plus BSC as the common comparator arm, to derive the indirect estimators of the efficacy of cabozantinib relative to regorafenib for the outcomes of interest. The method allows the randomisation of the RCTs to be preserved by utilising the relative treatment effects from each of the randomised trials. The main underlying assumption is that there is no difference in the distribution of effect modifying variables between trials, which allows the combination of their relative effects. The Bucher ITC for cabozantinib versus regorafenib included the OS primary, PFS secondary endpoints and safety of both trials (50).

Results – efficacy outcomes

The results of the Bucher ITC showed hazard ratios versus regorafenib that favoured cabozantinib for PFS [HR 0.96 (0.73, 1.26)] using RESORCE mRECIST criteria but favoured regorafenib using RECIST 1.1 criteria [HR 1.02 (0.78, 1.34)]. The results for OS favoured regorafenib [HR 1.23 (0.94, 1.61)], but the results were not statistically significant suggesting similar efficacy in terms of OS and PFS for both treatments.
The efficacy results are presented in Table 22.

Table 22: Summary of Bucher ITC results for cabozantinib plus BSC versus regorafenib plus BSC in ITT populations of their respective trials

Endpoint: relative effect measure	CELESTIAL: Cabozantinib versus placebo HR (95% CI)	RESORCE: Regorafenib versus placebo HR (95% CI)	Bucher ITC: Cabozantinib versus regorafenib HR (95% CI)
Overall survival	0.76 (0.63, 0.92)	0.62 (0.51, 0.75)	1.23 (0.94, 1.61)
Progression-free survival	0.44 (0.36, 0.52)	0.43 (0.35, 0.52)	1.02 (0.78, 1.34)

Abbreviations: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison.
Source: Abou-Alfa et al., 2018 (6), Exelixis, 2018 (44), Finn et al, 2018 (51), Waldschmidt et al, 2019 (52)

Log cumulative hazard plots and Schoenfeld residual plots were used to test the proportional hazards assumption underlying the Bucher ITC. Therefore, OS and PFS Kaplan-Meier curves from RESORCE were digitised and pseudo individual patient level data (IPD) generated, using the Guyot algorithm (53). The curves in the log cumulative hazard plot for OS were not parallel and cross (Figure 10). Furthermore,

the Schoenfeld residuals show correlation with time (Figure 11) and a Grambsch and Therneau test (a more formal statistical test based on the scaled Schoenfeld residuals) had a p-value of 0.0016. The findings suggested that the proportional hazards assumption was not satisfied for OS.

Figure 10: OS Log-log cumulative hazards

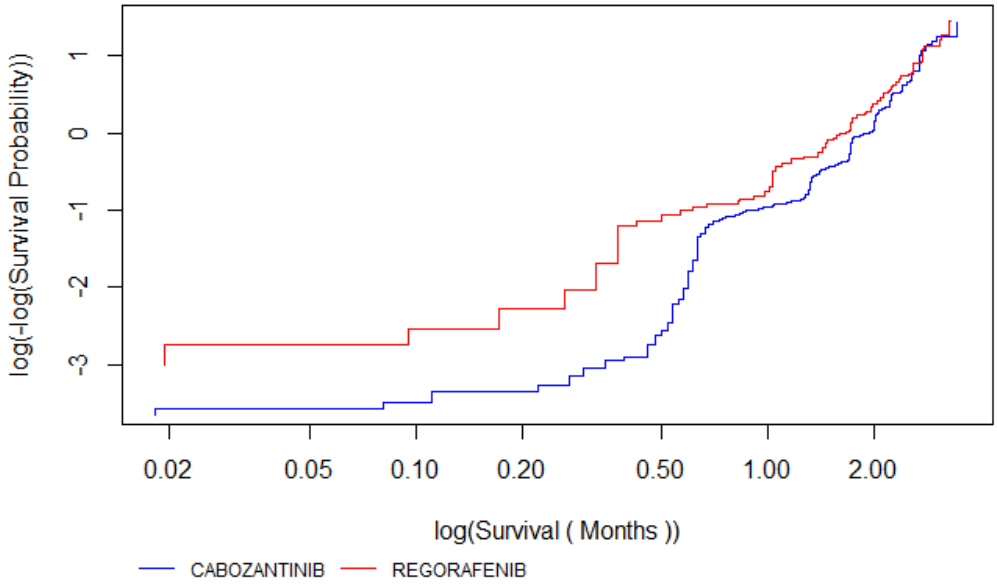
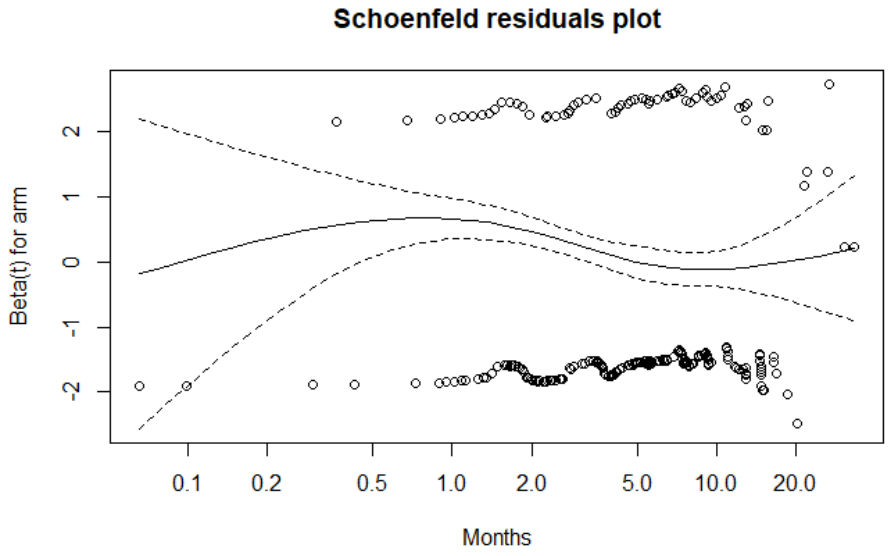


Figure 11: OS Schoenfeld residuals plot



However, for PFS, the curves in the log cumulative hazard plot were overlapping (Figure 12). The Schoenfeld residuals showed little correlation with time (Figure 13) and the Grambsch and Therneau test shows a p value of 0.73. This suggested that the proportional hazards assumption was satisfied for PFS.

Figure 12: PFS Log-log cumulative hazards

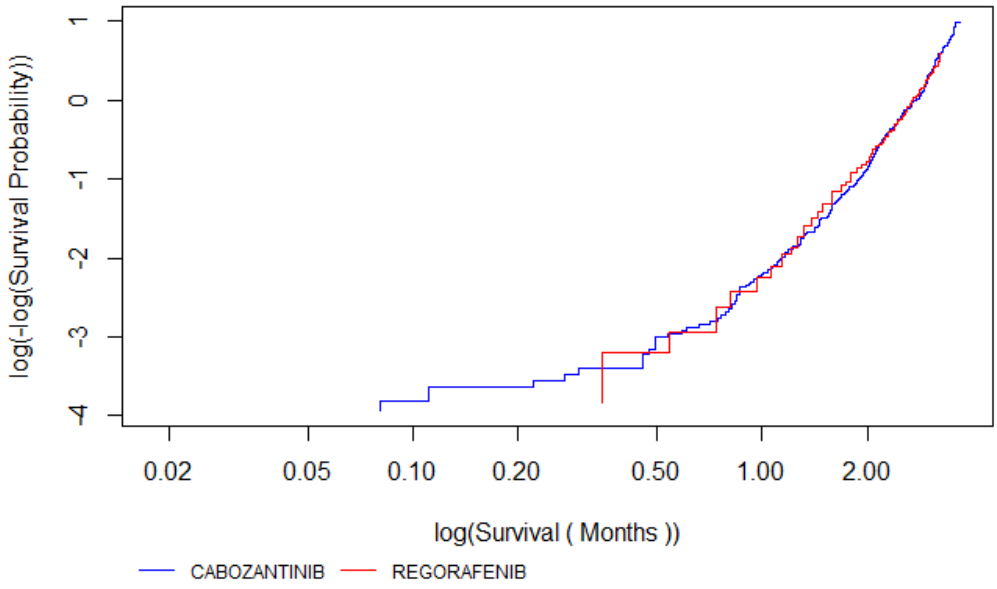
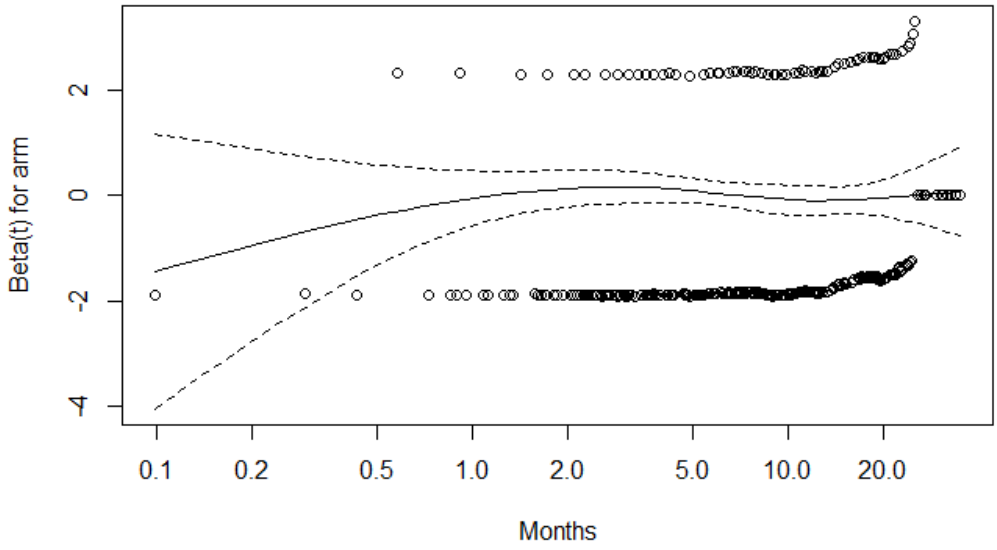


Figure 13: PFS Schoenfeld residuals plot



Results – safety outcomes

Treatment-emergent AEs with a grade 3/4 that occurred in ≥5% of patients in either arm was analysed. This is considered a standard approach as treatment-emergent grade 3/4 events are likely to be associated with higher costs and larger impact on quality of life than grade 1/2 events. This is consistent with previous submission to NICE in advanced HCC (1). Table 23 below presents the AEs considered in the analysis.

Table 23: Treatment-emergent AEs with a grade 3/4 that occurred in ≥5% of patients from CELESTIAL and RESORCE

Adverse Events	Cabozantinib n=467 n (%)	CELESTIAL placebo n=237 n (%)	Regorafenib n=374 n (%)	RESORCE placebo n=193 n (%)
Palmar-plantar erythrodysaesthesia syndrome	78 (16.7)	0 (0)	47 (12.6)	1 (0.5)
Hypertension	69 (14.8)	2 (0.8)	49 (13.1)	6 (3.1)
Elevated aspartate aminotransferase	36 (7.7)	11 (4.6)	19 (5.1)	10 (5.2)
Fatigue	39 (8.4)	6 (2.5)	24 (6.4)	3 (1.6)
Diarrhoea	42 (9.0)	2 (0.8)	9 (2.4)	0 (0)
Elevated bilirubin	0 (0)	0 (0)	25 (6.7)	4 (2.1)

Abbreviations: AEs, adverse events.
Source: Bruix et al,2017 (40), Exelixis, 2018 (44).

For the comparison of AEs, Bucher adjusted comparisons were only feasible when there were events in all arms of CELESTIAL and RESORCE. Therefore, only hypertension, elevated aspartate aminotransferase and fatigue AEs were compared. The results show no statistically significant differences between the AE ORs for cabozantinib and regorafenib. It should be noted that the small number of events results in large confidence intervals. The results are shown in Table 24.

Table 24: Summary of Bucher ITC safety results

Adverse Events	Regorafenib vs. Cabozantinib OR (95% CI)
Hypertension	0.2 (0.0-1.2)
Elevated aspartate aminotransferase	0.6 (0.2-1.6)
Fatigue	1.2 (0.3-5.6)

Abbreviations: CI, confidence interval; ITC, indirect treatment comparison; OR: odds ratio.

B.2.9.3 Matching-adjusted indirect comparison

Given the differences in baseline characteristics between CELESTIAL and RESORCE and the finding that the PH assumption may not be supported for OS, the efficacy of cabozantinib and regorafenib was compared using a MAIC as it provides a method of comparing absolute treatment effects while lowering the risk of bias associated with naïve unadjusted comparisons (54, 55).

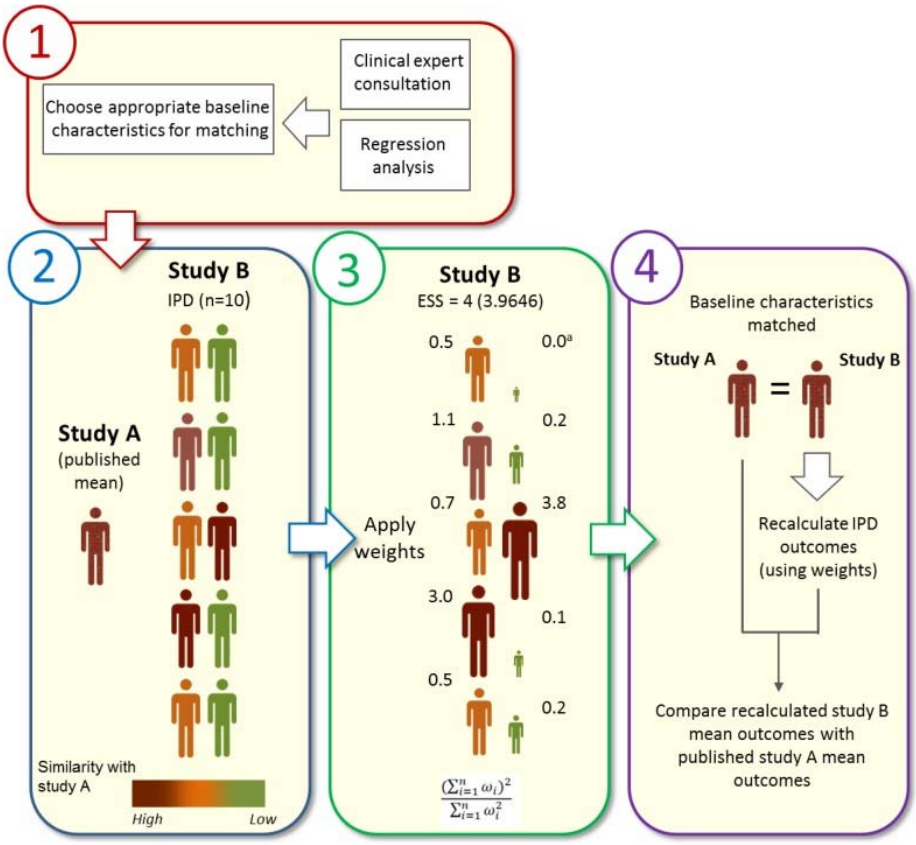
The MAIC analysis utilised a subpopulation from the CELESTIAL ITT population, specifically second-line hepatocellular carcinoma patients who had prior treatment with sorafenib (i.e., pure second-line patients) in order to compare to the RESORCE trial.

The method incorporates IPD, in this case available for CELESTIAL, which were reweighted to mimic the population of the RESORCE trial for which only aggregate results were available. The survival outcomes were recalculated for each pure second-line patient in CELESTIAL using the weighted data.

Methodology

An overview of the MAIC procedure is presented below in Figure 14.

Figure 14: Overview of MAIC procedure



Abbreviations: IPD, Individual patient data, ESS: Effective sample size.
Source: Nash et al, 2018 (56).

Baseline characteristics

Comparison of the patient characteristics of RESORCE with those of the pure second-line population of CELESTIAL suggest some differences remained in terms of ethnic mix, region, ECOG performance status and duration of prior sorafenib treatment between both populations. Other characteristics mentioned in Table 25 were similar or had minor differences. After removing subjects with missing values for the

characteristics, the pure second line CELESTIAL population was reduced from 495 to 484 patients (326 in the cabozantinib arm and 158 in the placebo arm).

Two different scenarios were considered to assess the impact of choosing different baseline characteristics for matching:

- In the first scenario (S1), which represents the base case, the baseline characteristics selected for matching were those deemed potential effect modifiers by the clinical experts.
- In the second scenario (S2), which serves as sensitivity check, the baseline characteristics selected for matching were those selected using the stepwise Akaike information criterion (AIC) regression strategy.

Reweighted baseline values of second-line subjects of CELESTIAL trial are presented below in Table 25.

Table 25. Comparison of reweighted baseline characteristics of subjects enrolled in CELESTIAL (pure 2nd line) and RESORCE

Treatment (N)	CELESTIAL pure 2nd line		RESORCE
	Pure 2nd line (S1)	Pure 2nd line (S2)	As reported
	Cabozantinib (N = 187.27)##	Cabozantinib (N = 303.24)##	Regorafenib (N = 374)
Age under 65 (%)	54.97##	53.34##	54.97##
Female (%)	18.63	12.04	12.04
Asia geographical region (%)	37.7	22.93	37.7
ECOG status 0 (%)	65.79	65.79	65.79
Child-Pugh class A (%)	97.91	98.86	97.91
Mean duration of sorafenib treatment (months)	11.63	7.52	11.63
Extrahepatic disease (%)	71.9	71.9	71.9
Macrovascular invasion (%)	28.62	28.62	28.62
Hepatitis B aetiology (%)	37.7	37.92	37.7
Alcohol use aetiology (%)	25.31	22.78	25.31
Hepatitis C aetiology (%)	20.77	24.53	20.77
AFP > 400ng/mL (%)	43.46	43.46	43.46
White (%)	35.95	58.15	35.95

Abbreviations: AFP; alpha fetoprotein; ECOG, Eastern Cooperative Oncology Group; N, number of patients; S, scenario; ESS; Effective Sample Size.
##Approximate ESS values.

Statistical analysis

The baseline characteristics used for the matching procedure were selected from the preliminary set based on their potential influence on key efficacy (PFS and OS) and safety outcomes (AEs).

The NICE Decision Support Unit Technical Support Document 18 recommends justifying the choice of matching parameters by clinical expert advice and empirical identification of all prognostic variables and effect modifiers included in the weighting model. The clinical relevance of potential matching variables was justified by clinical experts on a UK advisory board meeting on the 28th June 2018 and further validated at an advisory board meeting on 31st March 2021 (3, 57). The baseline characteristics available for matching in both trials and deemed potential effect modifiers by the clinical experts were age group, race, geographical region, ECOG performance status, Child-Pugh class, duration of prior sorafenib treatment, extrahepatic disease, macrovascular invasion, aetiology of HCC (Hepatitis B, alcohol use and Hepatitis C), and AFP level.

Proportions and means were published for RESORCE for the following characteristics and were available for CELESTIAL: age group, gender, geographical region, ECOG performance status, Child-Pugh class, duration of prior sorafenib treatment, extrahepatic disease, macrovascular invasion, extrahepatic disease and/or macrovascular invasion, aetiology of disease, AFP level and race. Patients recruited to the RESORCE trial had increased tolerability to sorafenib; however, due to lack of reported data, this variable could not be accounted for. Considering the limited data, duration of prior sorafenib was considered as a proxy for sorafenib tolerability.

All the aforementioned characteristics were presented as dichotomous variables for RESORCE (Bruix et al., 2017), except for duration of prior sorafenib treatment, aetiology of disease and race. Duration of prior sorafenib treatment is a continuous variable, aetiology of disease is a categorical variable with 6 categories (Hepatitis B, alcohol use, Hepatitis C, unknown, non-alcoholic steatohepatitis and other), and race is a categorical variable with 4 categories (white, Asian, black and other/not reported). Duration of prior sorafenib treatment has been reported as a mean (Finn et al., 2018),

and all the dichotomous and categorical variables have been reported as percentages. Aetiology and race have been dichotomised into multiple characteristics (one for each category).

The following measures were taken to remove characteristics suspected to introduce noise into the matching processor to be strongly correlated with other baseline covariates. For aetiology and race, categories including under 10% of subjects in each trial (such as non-alcoholic steatohepatitis aetiology and black race) or representing 'other' or 'not reported' were not matched. Asian race is very strongly correlated with Asia geographical region (e.g., all the patients enrolled in the CELESTIAL trial in Asia are of Asian race); it is therefore not matched. Similarly, 'extrahepatic disease and/or macrovascular invasion' is not matched as it is evidently very strongly correlated with each of its individual components. Matching the aforementioned characteristics would likely result in a loss of statistical power/efficiency and overmatched/overfitted data (if a covariate is already balanced across the two trials, except for random noise, matching it will just introduce additional noise into the system).

Additionally, effect modifiers for the primary survival endpoint, OS, are identified empirically via a stepwise AIC regression strategy. In this strategy, candidate baseline characteristics were added (or eliminated) from a regression model using a stepwise process based on the AIC. The stepwise model comparison was run in all directions (forward, backward and both) (58). In all cases, the predictors giving the lowest AIC were gender, ECOG performance status, extrahepatic disease, macrovascular invasion and AFP level. These predictors were clinically plausible effect modifiers, except for gender as per clinical feedback received from the advisory board and hence not included for matching (57). The baseline characteristics used for matching, and the matching scenarios considered are summarised in Table 26.

Table 26: Baseline characteristics selected for matching

Clinical expert selection (scenario 1)	Empirical analysis (scenario 2)
ECOG performance status	ECOG performance status
Baseline HCC disease per CRF (EHS and MVI)	Baseline HCC disease per CRF (EHS and MVI)
AFP level >400ng/ml	AFP level >400ng/ml
Age group	Gender
Child-Pugh class	
Duration of prior sorafenib treatment	

Clinical expert selection (scenario 1)	Empirical analysis (scenario 2)
Race	
Aetiology of HCC (Hepatitis B, alcohol use and Hepatitis C)	
Geographical region	

Abbreviations: AFP, alpha fetoprotein; CRF, case report form; ECOG, Eastern Cooperative Oncology Group; EHS, extrahepatic spread; HCC, hepatocellular carcinoma; MVI, macroscopic vascular invasion.

Safety

The estimated relative effects of cabozantinib versus placebo in the RESORCE population are found by taking weighted means of the AE outcomes in the CELESTIAL trial. These estimates have been generated using a linear model. This allows for the correct calculation of standard errors using a robust sandwich estimator (59). The log ORs of regorafenib versus placebo are computed using the reported data on AEs. The variance of the log ORs is approximated using the delta method. The indirect comparison estimates of cabozantinib versus regorafenib are constructed in the log OR scale, using the fact that they are equal to the estimated effects (log OR) of cabozantinib versus placebo minus the estimated effects of regorafenib versus placebo in the RESORCE population

Results

Rescaled weights

The distribution of the weights for Scenario 1 is examined in Figure 15, where the weights have been rescaled relative to the original unit weights of each individual. The histogram in Figure 16 examines the distribution of rescaled weights for Scenario 2. The histogram for Scenario 1 (Figure 15) shows that there are some very large, rescaled weights, with a maximum at 9.21. Scenario 2 reduces the presence of extreme weights (the maximum rescaled weight is 1.61), resulting in an approximate ESS which is very close to the original sample size and pulling the rescaled weights closer to one. Scenario 2 provides greater statistical power and precision than Scenario 1. However, Scenario 2 does not match some characteristics that are considered to be important effect modifiers by the clinical experts, and which differ considerably across trials (e.g. duration of prior sorafenib treatment and geographical region). Also, the automatic variable selection method employed only evaluates the most contributory predictor variables for the primary survival endpoint, OS, and not for

PFS or safety outcomes. In addition, the weighting of certain characteristics could drive the variables that have not been matched, moving the average for these variables further away from the values reported in RESORCE. However, this effect does not appear to be significant in the scenarios considered.

Figure 15: Histogram of rescaled weights (Scenario 1)

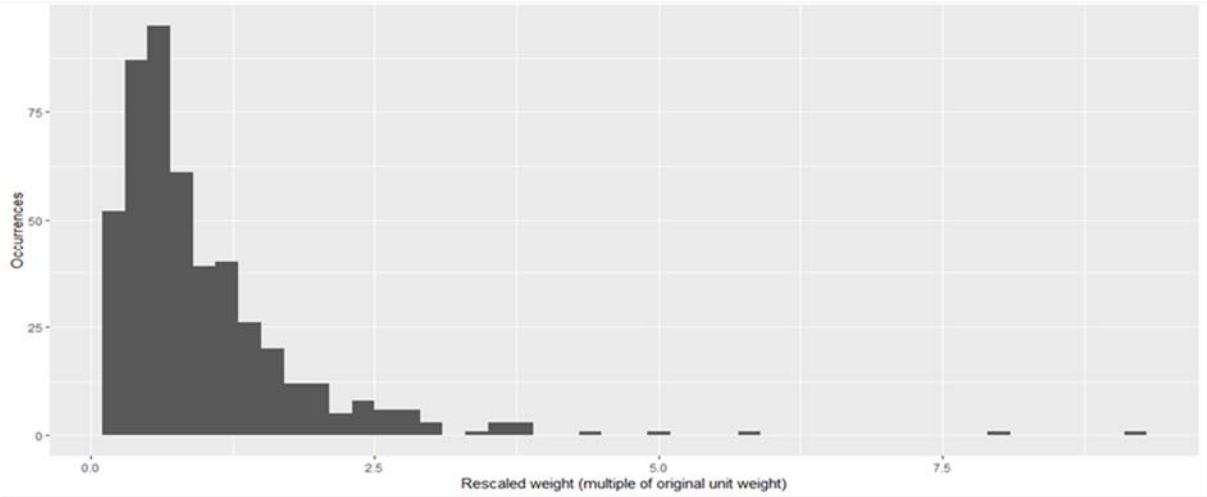
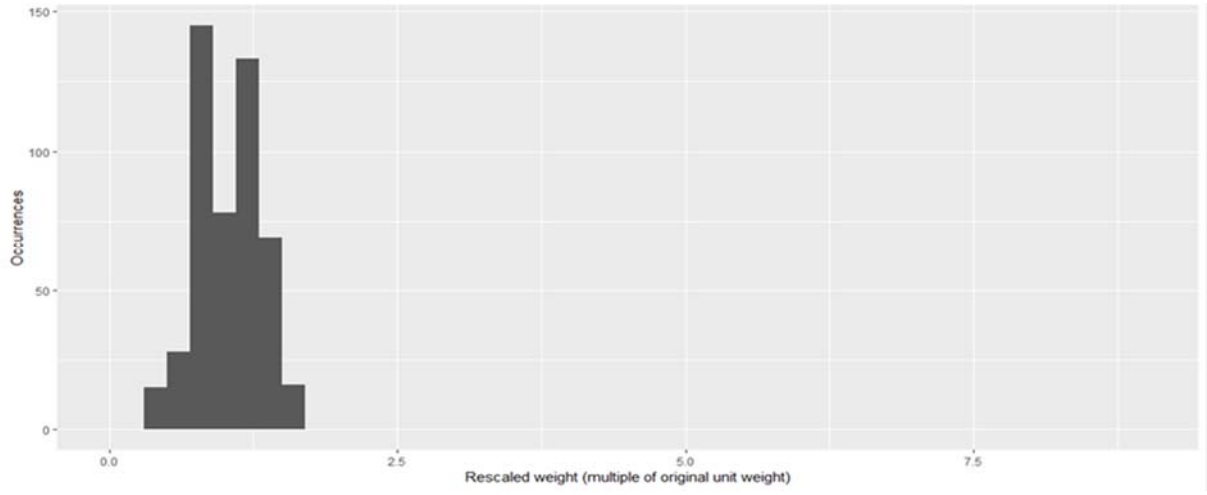


Figure 16: Histogram of rescaled weights (Scenario 2)



Efficacy outcomes

The selected PLD from CELESTIAL was adjusted to match aggregate data from RESORCE, survival outcomes were recalculated for each pure second-line patient in CELESTIAL using the weighted data. The pure second-line patient population from CELESTIAL had a median follow-up of 22.6 months. Table 27 presents summary

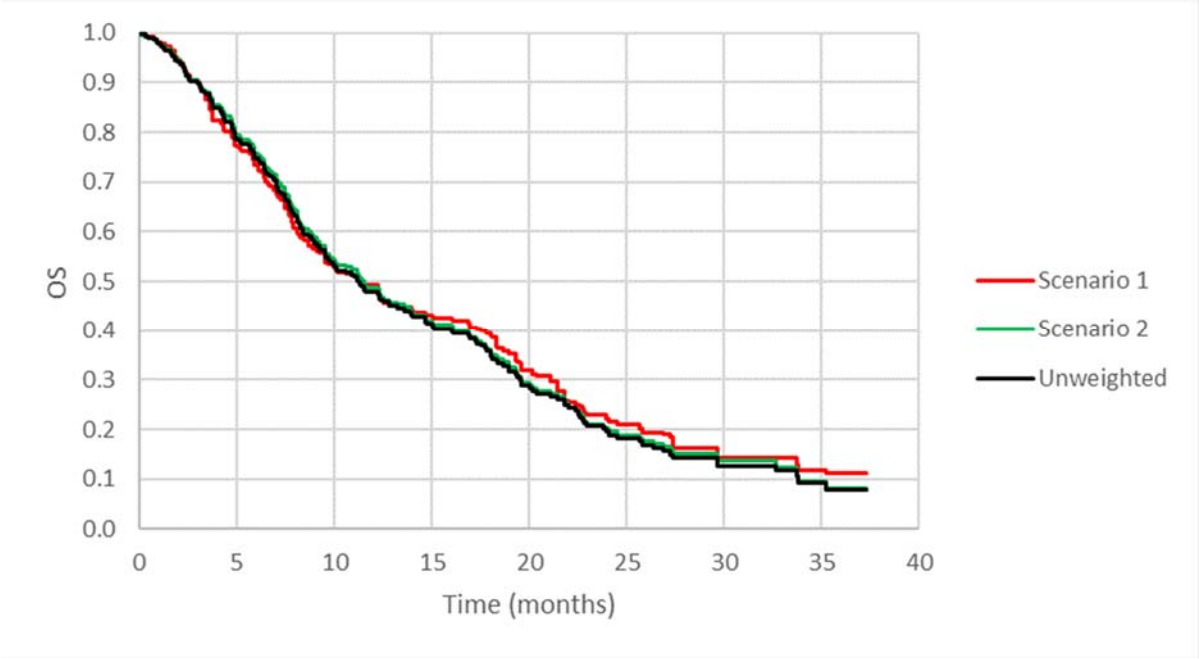
statistics with 95% confidence intervals for the (weighted and unweighted) Kaplan-Meier curves fitted to the cabozantinib and regorafenib survival data. For example, for regorafenib OS at the first quartile (i.e., 75% of patients are alive), 4.9 months have elapsed. Confidence intervals for quartiles use Woodruff's method: the interval is the intersection of the horizontal line at the specified quartile with the pointwise confidence band around the survival curve (60). This analysis suggests statistically significant differences at the 5% level for PFS but not for OS. Given the similarity between the scenarios, scenario 1 was considered the base case (60). A comparison of the cabozantinib weighted and unweighted scenarios are shown in Figure 17 and Figure 18 for OS and PFS respectively.

Table 27: Durations for endpoint Kaplan-Meier quartiles with 95% confidence intervals (in parentheses)

Treatment	PFS			OS		
	Q1 (months)	Q2 (months)	Q3 (months)	Q1 (months)	Q2 (months)	Q3 (months)
Regorafenib	1.45 (1.45-1.76)	3.19 (2.78-4.14)	6.99 (5.91-8.38)	4.90 (4.22-5.65)	10.79 (9.18-12.30)	20.96 (18.42-25.29)
Cabozantinib (unweighted pure-second line population)	2.07 (1.87-3.15)	5.52 (4.67-5.68)	9.20 (7.82-10.97)	5.91 (4.86-7.03)	11.24 (9.53-13.96)	21.85 (19.52-24.51)
Cabozantinib (weighted pure second-line population; Scenario 1)	2.37 (1.91-3.71)	5.59 (4.90-7.26)	9.56 (7.85-11.07)	5.78 (4.34-7.06)	11.37 (8.90-16.95)	22.74 (19.58-33.74)
Cabozantinib (weighted pure second-line population; Scenario 2)	2.10 (1.87-3.61)	5.55 (4.90-5.91)	9.20 (7.82-10.97)	6.21 (5.06-7.33)	11.50 (9.56-14.00)	22.05 (19.58-25.66)

Abbreviations: OS, overall survival; PFS, progression-free survival; Q, quartile.

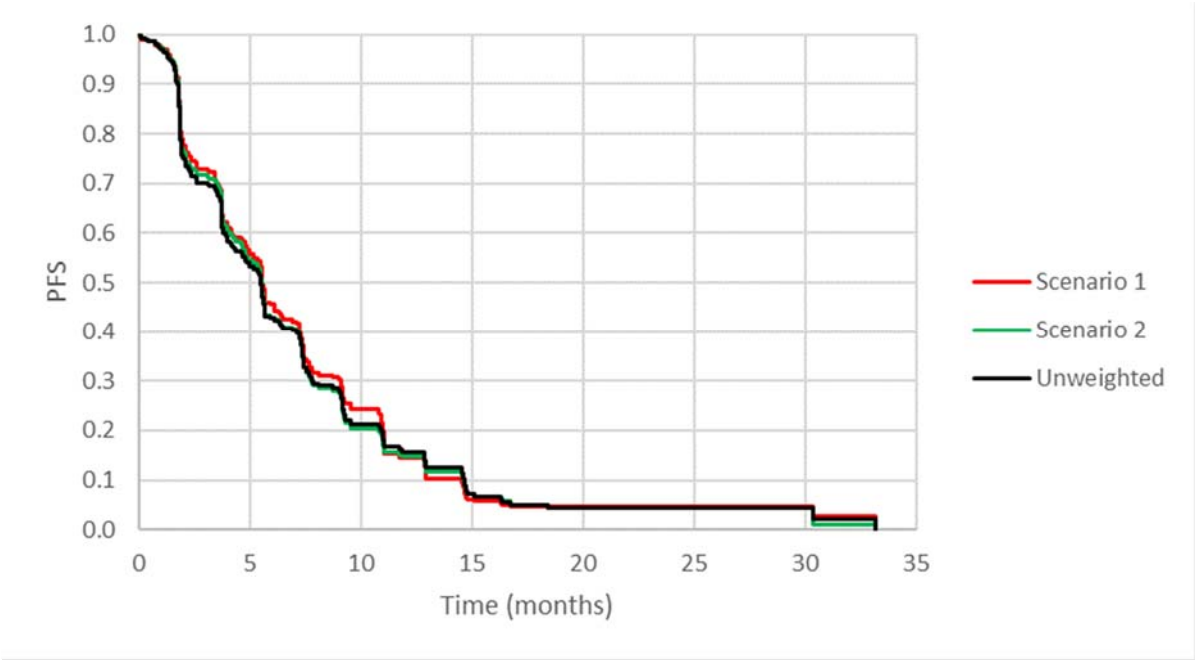
Figure 17: Weighted and unweighted cabozantinib OS KM



Abbreviations: KM, Kaplan-Meier; OS, overall survival

.

Figure 18: Weighted and unweighted cabozantinib PFS KM



Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival.

Safety outcomes

The log OR estimates are anchored because they use the common placebo arm. An anchored log OR estimate cannot be constructed for the diarrhoea AE because it has no occurrences in the placebo arm (giving a log OR of infinity for regorafenib versus placebo). For any AEs that do not occur in a given trial arm, approximate unanchored estimates of the log ORs are performed. Palmar-plantar erythrodysesthesia is another AE for which an unanchored estimate is performed, as it does not occur in the placebo arm of CELESTIAL pure second-line.

Table 28 presents the resulting anchored AE log ORs with 95% confidence intervals, standard errors and p-values.

Table 28: log ORs, confidence intervals, std. errors and p-values for treatment-emergent grade 3/4 AEs (cabozantinib vs. regorafenib)

Adverse event	CELESTIAL data	log OR	95% CI	standard error	p-value
Increased AST	Unweighted pure 2nd line	0.89	-0.31-2.09	0.61	0.1478
	Weighted pure 2nd line (S1)	0.79	-0.47-2.06	0.65	0.2201
	Weighted pure 2nd line (S2)	0.94	-0.29-2.17	0.63	0.1352
Elevated bilirubin	Unweighted pure 2nd line	-0.55	-3.01-1.91	1.25	0.6732
	Weighted pure 2nd line (S1)	-0.25	-2.73-2.23	1.26	0.8558
	Weighted pure 2nd line (S2)	-0.21	-2.67-2.25	1.26	0.8766
Fatigue	Unweighted pure 2nd line	0.07	-1.65-1.79	0.88	0.9404
	Weighted pure 2nd line (S1)	0.09	-1.77-1.94	0.95	0.9313
	Weighted pure 2nd line (S2)	0.4	-1.35-2.14	0.89	0.671
Hypertension	Unweighted pure 2nd line	1.73	-0.45-3.91	1.11	0.1207
	Weighted pure 2nd line (S1)	2.1	-0.1-4.3	1.12	0.0611
	Weighted pure 2nd line (S2)	1.72	-0.47-3.9	1.11	0.1239
Diarrhoea	Unweighted pure 2nd line	1.55	0.8-2.3	0.38	0.0001

Adverse event	CELESTIAL data	log OR	95% CI	standard error	p-value
(unanchored)	Weighted pure 2nd line (S1)	1.74	1-2.48	0.38	<0.0001
	Weighted pure 2nd line (S2)	1.68	0.94-2.43	0.38	<0.0001
Palmar-plantar erythrodysesthesia (unanchored)	Unweighted pure 2nd line	0.3	-0.17-0.77	0.24	0.2103
	Weighted pure 2nd line (S1)	0.05	-0.4-0.5	0.23	0.848
	Weighted pure 2nd line (S2)	0.3	-0.15-0.76	0.23	0.1934

Abbreviations: AE; adverse event; CI, confidence interval; OR, odd ratio; S, scenario.

Sensitivity analysis of the anchored MAIC

In order to assess differences between cabozantinib and regorafenib OS and PFS, the proportional hazards assumption was assessed. Figure 19 presents the log-cumulative hazard plot of weighted cabozantinib (Scenario 1) versus regorafenib for the PFS outcome. The curves remain parallel till after month 10 where the curves eventually cross. This would suggest that the proportional hazards assumption is not satisfied for the PFS outcome however there are low patient numbers generating the tail of these curves. The plot of the scaled Schoenfeld residuals (Figure 20) shows a degree of flatness however the Grambsch-Therneau test has a p-value of 0.0002 which indicates a non-zero slope.

Figure 19: PFS log-cumulative hazard plot for weighted cabozantinib (Scenario 1) versus regorafenib

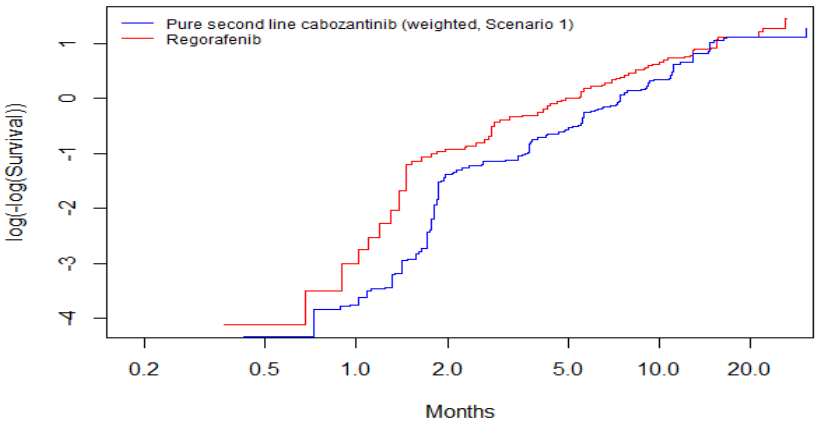


Figure 20: Scaled Schoenfeld residuals for PFS for weighted (Scenario 1) pure second-line cabozantinib versus regorafenib

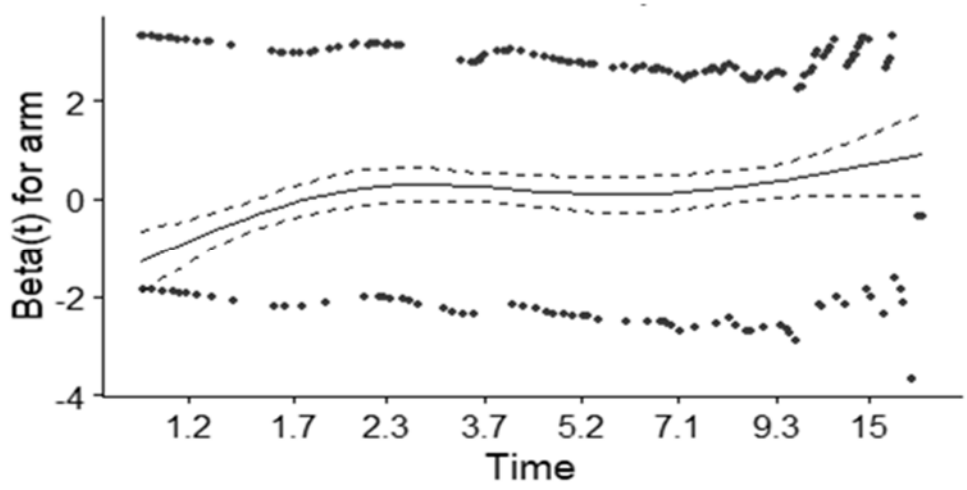


Figure 21 presents the corresponding log-cumulative hazard plot for the OS outcome. The OS curves cross at several instances. These intertwined curves suggest that the OS outcomes of the groups are similar. Similar to PFS, the plot of the scaled Schoenfeld residuals (Figure 22) shows a degree of flatness however the Grambsch-Therneau test (p-value 0.0029) indicates a non-zero slope as well.

Figure 21: OS log-cumulative hazard plot for weighted cabozantinib (Scenario 1) versus regorafenib

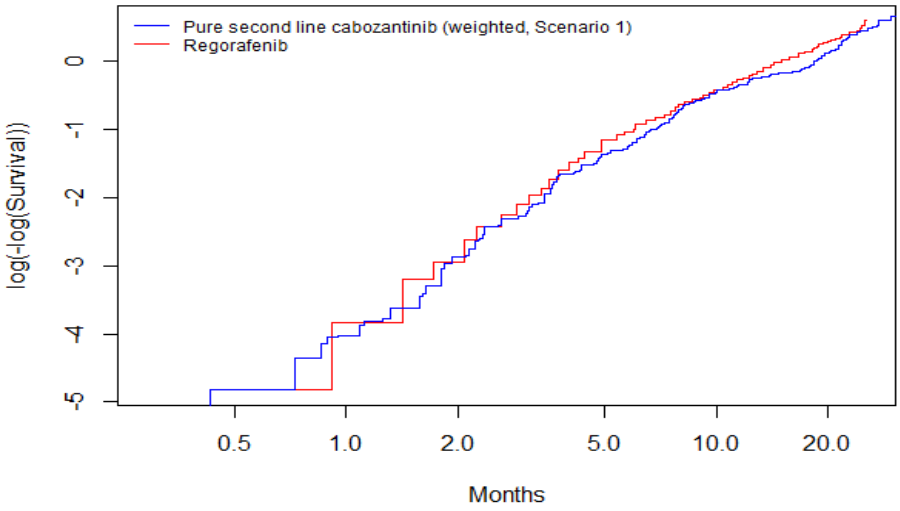
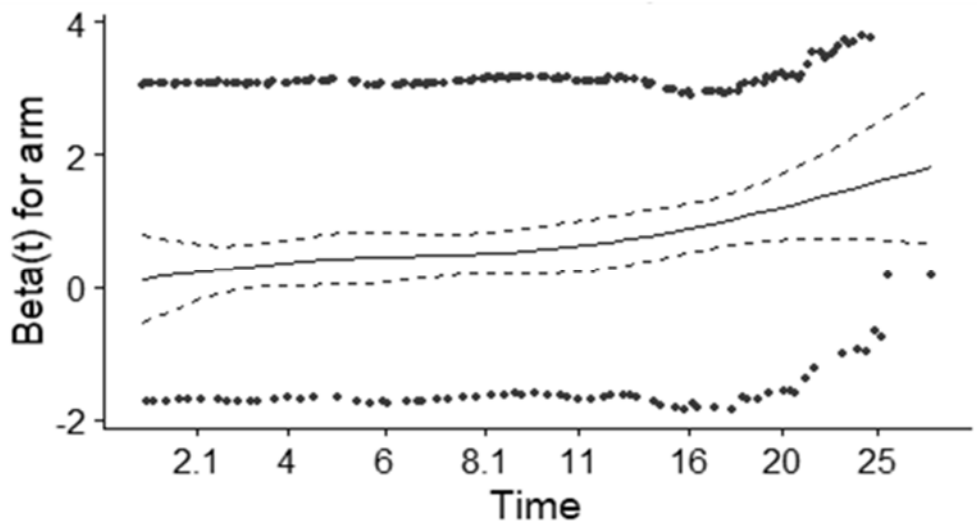


Figure 22: Scaled Schoenfeld residuals for PFS for weighted (Scenario 1) pure second-line cabozantinib versus regorafenib



Given the uncertainty of the proportional hazard assumption for both endpoints, a range of models were explored which would further assess the uncertainty of whether there was any difference in treatment effect between cabozantinib and regorafenib, as summarised below:

- An anchored analysis assuming that the proportional hazards assumption holds between cabozantinib and regorafenib. This analysis uses a constant HR of weighted CELESTIAL data and RESORCE to generate a hazard ratio between cabozantinib and regorafenib;
- An anchored analysis assuming that the proportional hazards assumption does not hold. This analysis explores if there is any difference in treatment effect emerging between cabozantinib and regorafenib over time. This is conducted by generating time-varying hazard ratios from hazard profiles of fitted parametric models to the weighted CELESTIAL and RESORCE data;
- An unanchored analysis comparing the treatment effect by using fitted parametric models to weighted cabozantinib and regorafenib data.

Anchored analysis using constant HR

The results of an anchored comparison between cabozantinib and regorafenib using a constant hazard ratio are shown in Table 29. The hazard ratio of cabozantinib versus

regorafenib shows a point estimate that favours PFS for cabozantinib, while the opposite for OS. Both of these results are not statistically significant.

Table 29. Results of anchored comparison using a constant hazard ratio

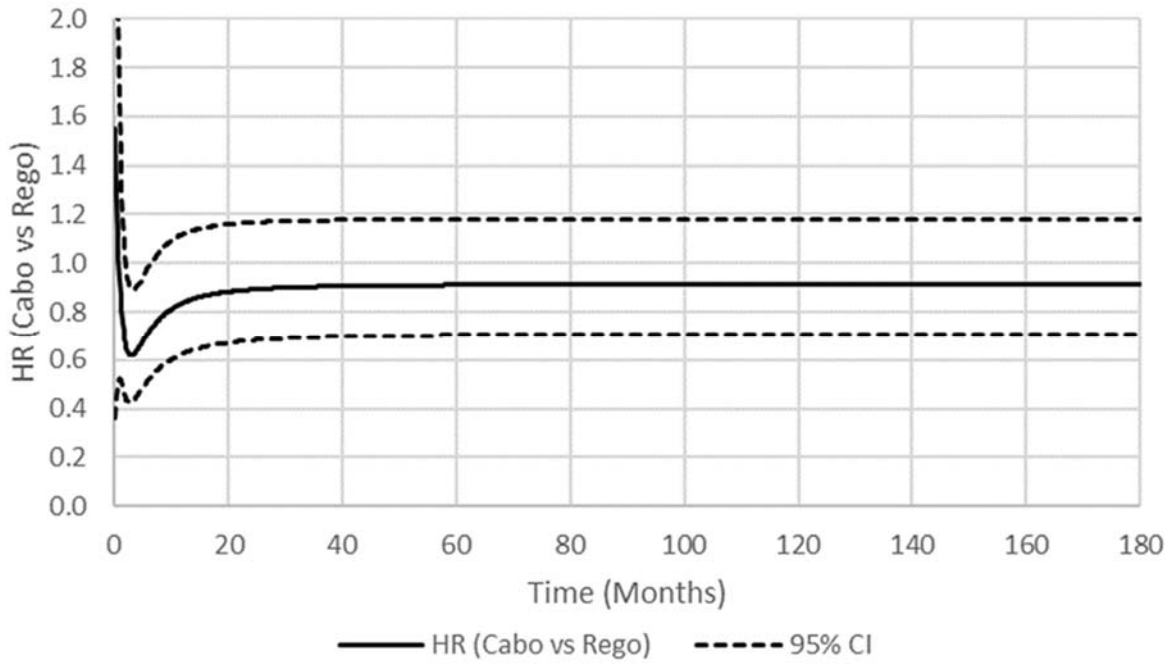
Endpoint: relative effect measure	Weighted CELESTIAL: Cabozantinib versus placebo HR (95% CI)	RESORCE: Regorafenib versus placebo HR (95% CI)	Cabozantinib versus regorafenib HR (95% CI)
Overall survival	0.73 (0.54, 0.99)	0.63 (0.50, 0.79)	1.15 (0.79, 1.69)
Progression-free survival	0.36 (0.28, 0.48)	0.46 (0.37, 0.56)	0.79 (0.56, 1.11)

Abbreviations: CI, confidence interval; HR, hazard ratio.

Anchored analysis using time-varying HR

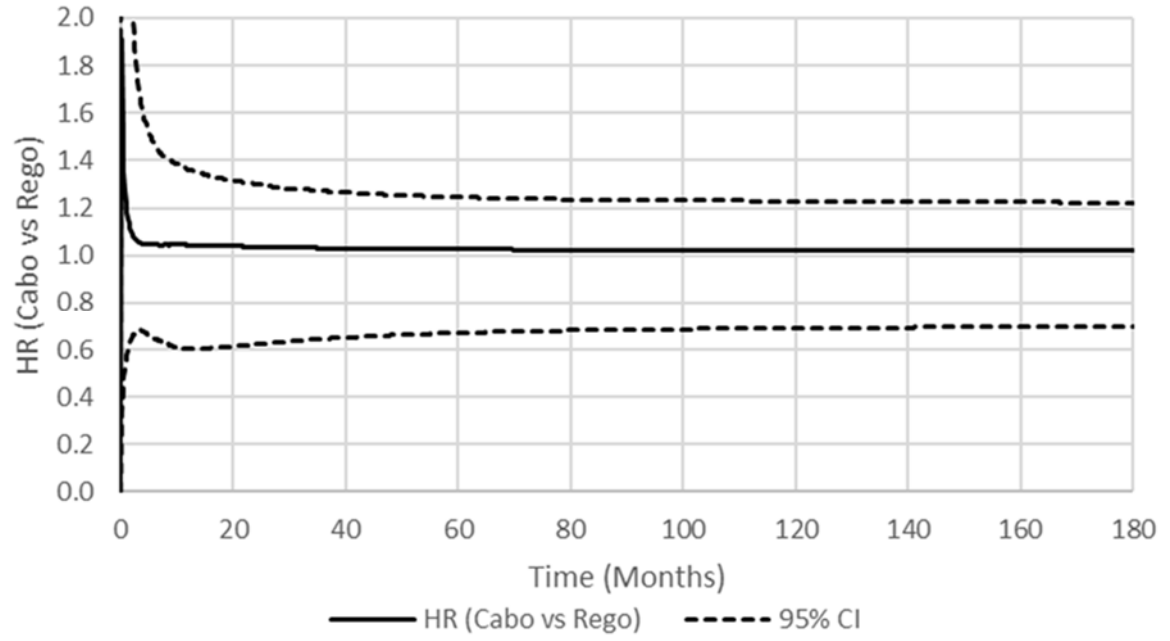
The result of the anchored analysis using time-varying hazard ratios generated from the log-logistic model is shown in Figure 23 for PFS and in Figure 24 for OS. For both endpoints the log-logistic model was the best fitting by AIC and Bayesian information criterion (BIC); however, the other standard parametric models were tested, and the results are shown in Appendix L. The results across the models show that over time, the hazard ratio is not statistically different from 1, indicating no difference in treatment effect. Furthermore, the hazard ratio is generally seen to be constant and near 1 as the treatment effect is extrapolated which suggests equivalence in treatment effect over time. Similar to the constant hazard ratio analysis, the point estimate shows conflicting direction of treatment benefit as there is a benefit for cabozantinib for PFS but a benefit for regorafenib for OS.

Figure 23: Log-logistic model for time-varying hazard ratio of cabozantinib versus regorafenib for progression-free survival endpoint



Abbreviations: Cabo, cabozantinib; CI, Confidence interval; HR, Hazard ratio; Rego, regorafenib.

Figure 24: Log-logistic model for time-varying hazard ratio of cabozantinib versus regorafenib for overall survival endpoint

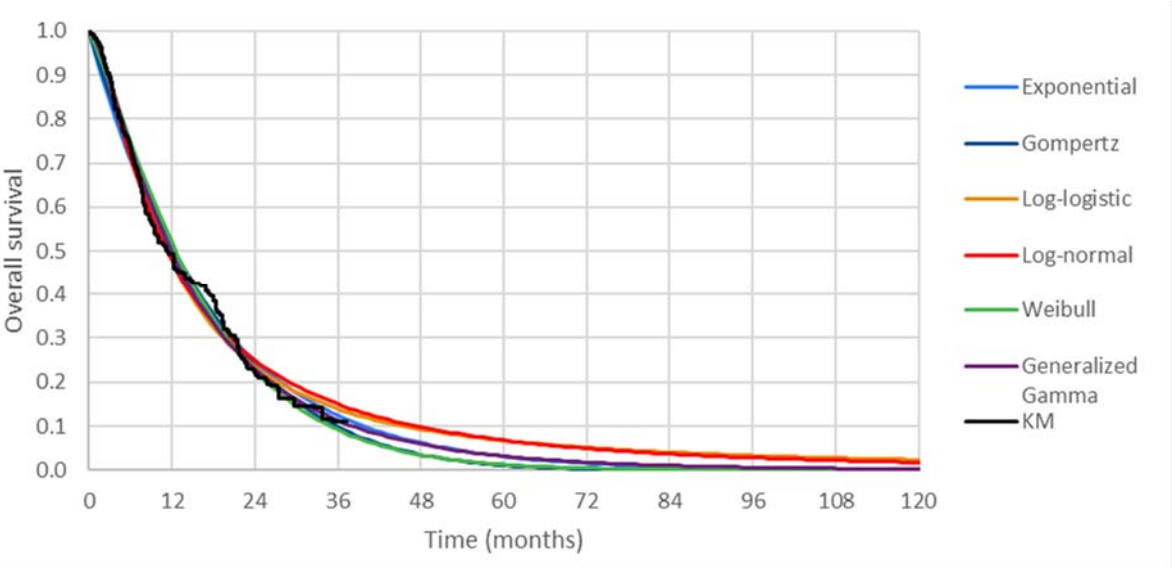


Abbreviations: Cabo, cabozantinib; CI, confidence interval; HR, hazard ratio; Rego, regorafenib.

Unanchored analysis using independent parametric models

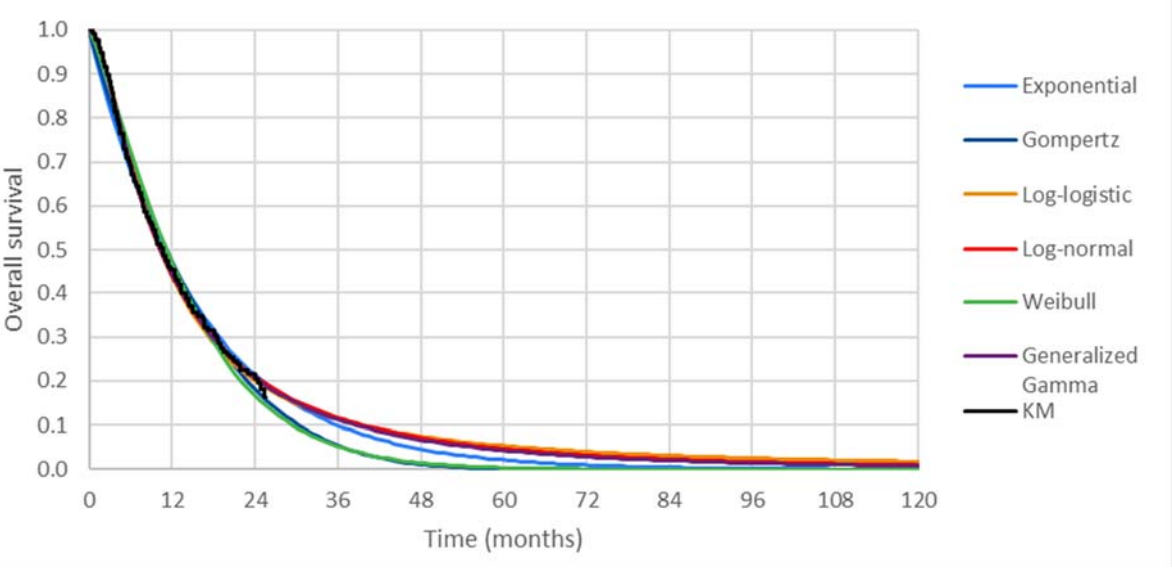
The parametric fits for the weighted cabozantinib OS data are shown in Figure 25, and the parametric fits for the regorafenib OS from the RESORCE trial are shown in Figure 26. The AIC and BIC estimates are shown in Table 30.

Figure 25: Parametric fits for weighted cabozantinib OS



Abbreviations: KM, Kaplan-Meier; OS, overall survival

Figure 26: Parametric fits for regorafenib OS from the RESORCE trial



Abbreviations: KM, Kaplan-Meier; OS, overall survival

Table 30. AIC and BIC statistics for weighted cabozantinib and regorafenib OS parametric fits

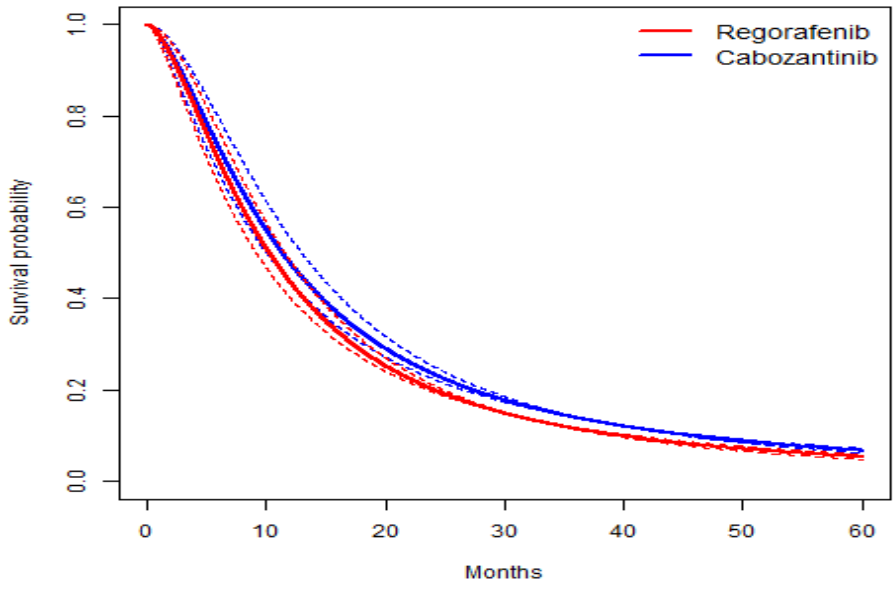
Endpoint / Model	Weighted CELESTIAL		RESORCE	
	AIC	BIC	AIC	BIC
Exponential	1678.56	1682.34	1740.62	1744.56

Endpoint / Model	Weighted CELESTIAL		RESORCE	
	AIC	BIC	AIC	BIC
Weibull	1672.09	1679.67	1727.96	1735.84
Gompertz	1678.39	1685.96	1739.24	1747.11
log-logistic	1668.20	1675.78	1716.81	1724.68
log-normal	1675.18	1682.75	1712.17	1720.05
Generalised gamma	1668.37	1679.74	1714.10	1725.92

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion.; OS, overall survival

The results of the unanchored analysis for OS are shown in Figure 27 using a log-logistic model. Confidence intervals were produced by simulating a large bootstrap-like sample from the asymptotic normal distribution of the maximum likelihood estimates of the parameters (61). In total, 100,000 random samples were drawn to ensure that the recovered mean and median survival times were stable to two decimal places through different runs. The OS curves show a large amount of overlap until year 1 when cabozantinib begins to show a relatively small benefit over regorafenib. Cabozantinib has a larger point estimate for mean OS (24.65 vs. 21.17 months) and a higher median OS (11.40 versus 10.29 months).

Figure 27: Unanchored results for OS

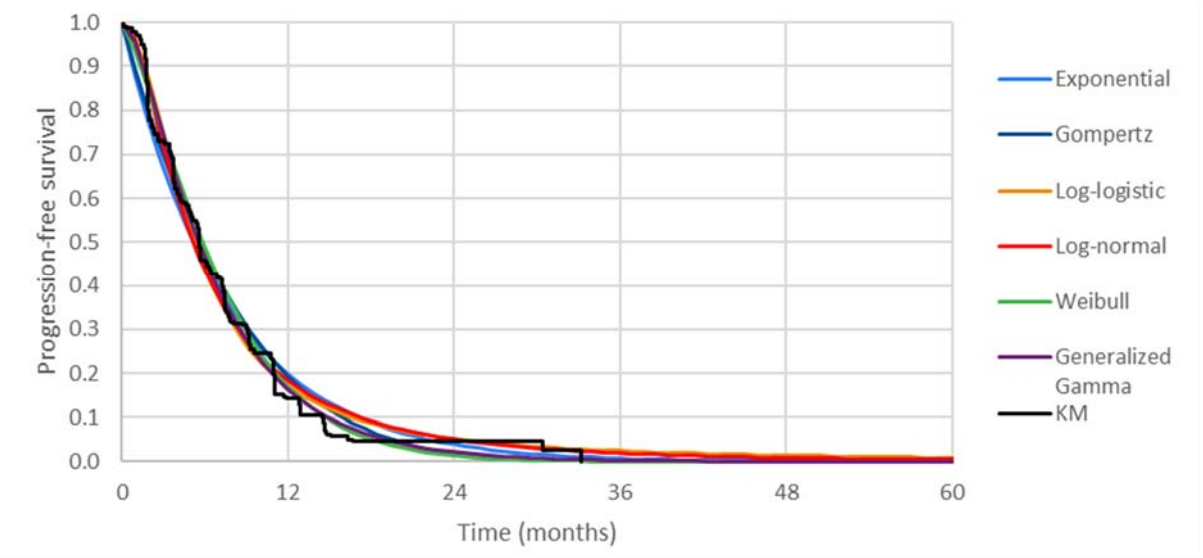


Abbreviations: OS, overall survival.

The parametric fits for the weighted cabozantinib PFS data are shown in Figure 28 and the parametric fits for the regorafenib PFS from the RESORCE trial are shown in

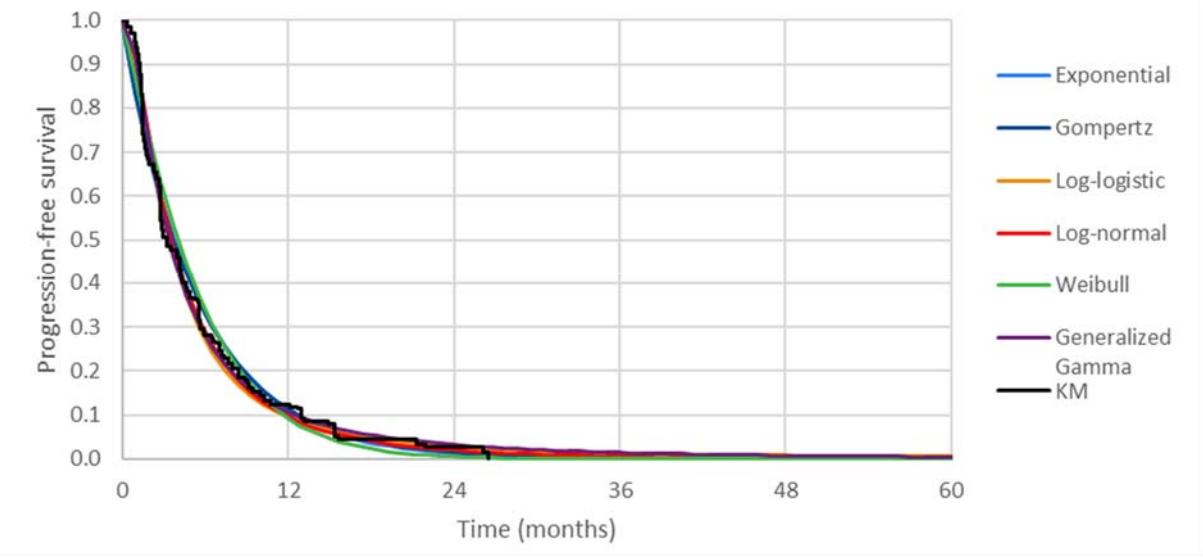
Figure 29. The AIC and BIC statistics are shown in Table 31. The generalised gamma is selected as the base case model due to the better statistical fit.

Figure 28: Parametric fits for weighted cabozantinib PFS



Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival

Figure 29: Parametric fits for regorafenib PFS from the RESORCE trial



Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival

Table 31. AIC and BIC statistics for weighted cabozantinib and regorafenib PFS parametric fits

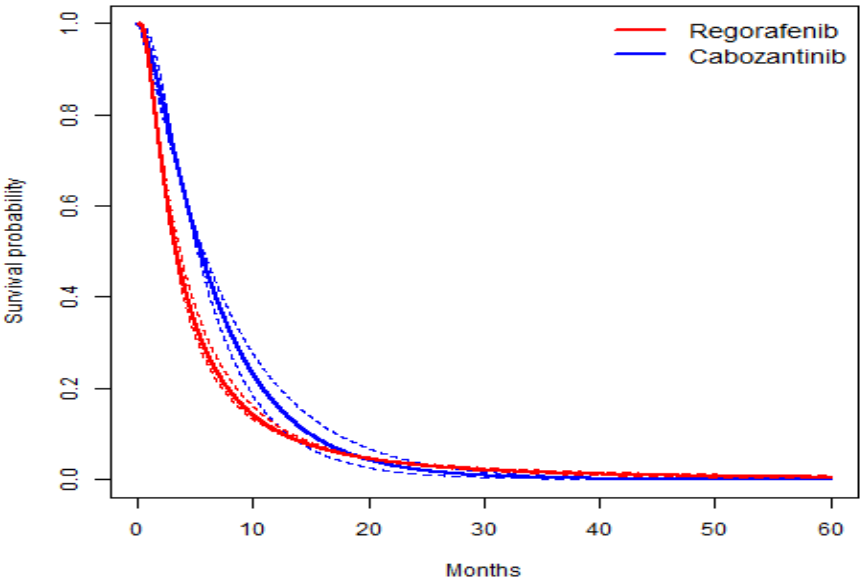
Endpoint / Model	Weighted CELESTIAL		RESORCE	
	AIC	BIC	AIC	BIC
Exponential	1480.30	1484.09	1641.66	1645.60

Endpoint / Model	Weighted CELESTIAL		RESORCE	
	AIC	BIC	AIC	BIC
Weibull	1457.16	1464.73	1634.92	1642.79
Gompertz	1476.18	1483.75	1643.38	1651.26
log-logistic	1453.83	1461.41	1590.28	1598.15
log-normal	1467.01	1474.58	1577.40	1585.27
Generalised gamma	1450.61	1461.97	1575.13	1586.94

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion.; PFS, progression-free survival

The results of the unanchored analysis for PFS are shown in Figure 30 using a generalised gamma model. The models show a statistically significant benefit for cabozantinib until approximately 1 year when the PFS curves show little difference for the rest of the time horizon. Cabozantinib has a larger point estimate for mean PFS than regorafenib (7.17 vs. 6.04 months) and higher median PFS (5.49 vs. 3.39).

Figure 30: Unanchored results for PFS



Abbreviations: PFS, progression-free survival.

B.2.9.4 Discussion and conclusions of indirect treatment comparisons

There was presence of between-study heterogeneity among CELESTIAL and RESORCE trials, namely the increased tolerability of patients to sorafenib in RESORCE and the inclusion of third-line patients in the CELESTIAL population. Despite the different populations, the Bucher approach showed that for the point estimates, OS favoured regorafenib but PFS slightly favoured cabozantinib. None of these results are statistically significant. The proportional hazards assumption did not

hold for OS; therefore, the treatment effect may not be representative as a constant hazard ratio.

When adjusting for population differences through the MAIC, the anchored analysis showed that cabozantinib has a higher point estimate than regorafenib for PFS; however, regorafenib was associated with higher OS (point estimate) than cabozantinib. None of these results were statistically significant. Relaxing the proportional hazards assumption through the time-varying hazard ratio analysis showed no significant difference for the treatment effect over time. The unanchored MAIC as a scenario analysis to the anchored approach showed that cabozantinib may achieve a similar OS and prolonged PFS compared with regorafenib. The improvement in PFS was statistically significant in favour of cabozantinib

A previously published MAIC study using real world evidence (RWE) for regorafenib showed similar results to that provided in this submission. The Casadei Gardini et al. analysis used data from 278 patients who received regorafenib as a second-line therapy after previous treatment with sorafenib for unresectable HCC. This group of patients also included those intolerant to sorafenib as well as tolerant, whereas the RESORCE trial only had sorafenib tolerant patients. Published aggregate data for the subgroup of CELESTIAL patients who received sorafenib as the only prior therapy were used in the analysis for cabozantinib data (62). This methodology estimates the effect of the regorafenib treatment in the patient population that received cabozantinib. The results found cabozantinib to have a statistically significant benefit over regorafenib in terms of PFS in all prior sorafenib patient populations [HR 0.50 (0.41-0.62)]. It also found a benefit in terms of OS with point estimates in favour of cabozantinib versus regorafenib [HR 0.83 (0.62-1.09)] but this was not statistically significant (63). Other network meta-analyses (NMAs) that have been conducted and reported in the literature have similarly found no statistically significant difference between the two treatment options in terms of survival or safety endpoints. The OS and PFS results are summarised in Table 32.

Table 32. Results from ITCs conducted in the literature

Study	Overall survival (HR 95% CI)	Progression-free survival (HR 95% CI)
Wang et al.2020 (64)	Rego vs Cabo: 0.82 (0.63-1.1)	Rego vs Cabo: 1.1 (0.80-1.4)
Bakouny et al. 2018 (65)	Rego vs Evero: 0.60 (0.44-0.51) Cabo vs Evero: 0.72 (0.55-0.95)	Rego vs Evero: 0.46 (0.35-0.62) Cabo vs Evero: 0.47 (0.36-0.63)
Sonbol et al. 2020 (66)	Rego vs Cabo: 0.82 (0.62-1.07)	Rego vs Cabo: 1.04 (0.79-1.36)
Park et al. 2021 (67)	Cabo vs Rego: 0.96 (0.54-1.68)	-
Casadei Gardini et al. 2021 (63)	Cabo vs Rego: 0.83 (0.62-1.09) Subgroups: Prior sorafenib < 3 months: Cabo vs Rego: 0.68 (0.39-1.16) Prior sorafenib 3 to 6 months: Cabo vs Rego: 0.66 (0.42-1.02) Prior sorafenib > 6 months: Rego vs Cabo: 0.89 (0.52-1.51)	Cabo vs Rego: 0.50 (0.41-0.62) Subgroups: Prior sorafenib < 3 months: Cabo vs Rego: 0.33 (0.21-0.50) Prior sorafenib 3 to 6 months: Cabo vs Rego: 0.53 (0.37-0.75) Prior sorafenib > 6 months: Cabo vs Rego: 0.60 (0.38-0.94)

Abbreviations: Cabo: cabozantinib; Evero, everolimus; HR, hazard ratio; ITC, indirect treatment comparison; Rego: regorafenib.

The AE analysis using a Bucher approach, showed different point estimates for AEs that were able to be analysed through the Bucher approach, but the results were not significant. When using the MAIC methodology, only diarrhoea shows statistically significant differences at the 5% level. However, this estimate is unreliable because the grade 3/4 treatment-emergent AE only occurs twice for the CELESTIAL placebo arm and never occurs for the RESORCE placebo arm. The patients in RESORCE were tolerant to sorafenib and this would reduce the occurrence of grade 3/4 treatment-emergent diarrhoea. Some of the anchored log ORs are very large (e.g., the estimates for hypertension are close to 2), probably a result arising from very small counts in the data, particularly in the CELESTIAL placebo arm, which make the estimates unprecise and drive them upward.

The RWE data shows that cabozantinib has a similar toxicity profile to that observed in the CELESTIAL trial with certain grade 3+ AEs of interest occurring closer to that of the numbers reported in the RESORCE trial (68, 69).

In conclusion, the ITC results suggest that cabozantinib has comparable or greater clinical efficacy and similar tolerability compared to regorafenib in the context of a RTK inhibitor.

B.2.9.5 Uncertainties in the indirect and mixed treatment comparisons

The population differences between the trials introduced bias into the Bucher analysis. Therefore, a MAIC was conducted to reduce the impact of these variables on the results. The effective sample size for the MAIC remained large with 265.53 for scenario 1 and 452.31 for scenario 2. There were some large, rescaled weights in scenario 1 with a maximum of 9.21 but scenario 1 matched with more characteristics that are considered to be important effect modifiers by the clinical experts, and which differ considerably across trials (e.g. duration of prior sorafenib treatment and geographical region). The two scenarios produced similar results.

A negative outcome control was conducted as a form of validation. This compared the weighted placebo arm of CELESTIAL and the placebo arm of RESORCE. The MAIC can balance observed patient characteristics but there is still the potential for residual confounding due to unobserved differences between trials. The recovered HR for OS (CELESTIAL placebo vs. RESORCE placebo) was 0.87 (95% confidence interval 0.67-1.15; p-value 0.326). For Scenario 2, the estimated HR for OS is 0.88 (95% confidence interval 0.68-1.14; p-value 0.326). In both cases, the HR was close to one. The recovered HR for PFS was 0.69 (95% confidence interval 0.55-0.87; p-value 0.00158). For Scenario 2, the estimated HR for PFS was 0.72 (95% confidence interval 0.58-0.90; p-value 0.00328). This would suggest that, even after matching, there remains important cross-trial differences in the placebo arms. There is therefore some sort of residual imbalance impacting the PFS outcomes. This adds uncertainty to any superiority claim in terms of PFS benefit for cabozantinib over regorafenib and thus no superiority is assumed in this submission as a conservative assumption.

The uncertainty regarding the proportion hazards assumption was explored by investigating the trend of the hazard ratio over time between cabozantinib and regorafenib. The time-varying hazard ratio analysis was able to show that there was no significant difference for the treatment effect over time. A further sensitivity was conducted by not using the hazard ratio to represent the treatment effect but instead

fit independent curves to the cabozantinib and regorafenib arms. This showed similar or better treatment effect for cabozantinib which is in line with the conservative assumption of no superiority between treatments.

B.2.10. Adverse reactions

B.2.10.1 Summary of safety data

In the CELESTIAL trial, the population for the analysis of safety (safety population) comprised of all patients who received at least one dose of study drug (n=704; n=467 for cabozantinib and n=237 for placebo). In the safety population in the CELESTIAL trial, patients in the placebo group received a mean (±standard deviation) daily dose of 52.85 mg (±11.1) and those in the cabozantinib group received a mean daily dose of 36.56 mg (±13.8) (44).

Cabozantinib was generally well tolerated. AEs frequently reported with cabozantinib were typical of those with VEGFR-TKI therapies. An overview of safety data from the CELESTIAL trial is provided in Table 33 and Table 34.

Table 33: The CELESTIAL trial: summary of safety data (safety population)

Adverse Events	Cabozantinib n=467 n (%)	Placebo n=237 n (%)
Any AE (all grades)	460 (99)	219 (92)
Grade 3 or 4 AEs	316 (68)	86 (36)
Treatment-related AEs	439 (94)	148 (62)
SAEs	232 (50)	87 (37)
Treatment-related SAEs	82 (18)	14 (5.9)
Treatment-related Grade 5 AE	6 (1.3)	1 (0.4)
Deaths (at any time, excluding PD)	314 (67)	167 (70)
AE leading to dose modification	416 (89)	94 (40)
AE leading to discontinuation of study drug	96 (21)	10 (4.2)

Abbreviations: AEs, adverse events; PD, progressive disease; SAEs, serious adverse events.
Source: Abou-Alfa et al., 2018 (6), Exelixis, 2018 (44)

Table 34. AEs* (any grade) reported in ≥10% of patients in either treatment group

Event	Cabozantinib (number of patients (percent)			Placebo (number of patients (percent)		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
Any AE	460 (99)	270 (58)	46 (10)	219 (92)	80 (34)	6 (3)
Diarrhoea	251 (54)	45 (10)	1 (<1)	44 (19)	4 (2)	0
Decreased appetite	225 (48)	27 (6)	0	43 (18)	1 (<1)	0

Event	Cabozantinib (number of patients (percent)			Placebo (number of patients (percent)		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
PPES	217 (46)	79 (17)	0	12 (5)	0	0
Fatigue	212 (45)	49 (10)	0	70 (30)	10 (4)	0
Nausea	147 (31)	10 (2)	0	42 (18)	4 (2)	0
Hypertension	137 (29)	73 (16)	1 (<1)	14 (6)	4 (2)	0
Vomiting	121 (26)	2 (<1)	0	28 (12)	6 (3)	0
Increase in AST level	105 (22)	51 (11)	4 (1)	27 (11)	15 (6)	1 (<1)
Asthenia	102 (22)	31 (7)	1 (<1)	18 (8)	4 (2)	0
Dysphonia	90 (19)	3 (1)	0	5 (2)	0	0
Constipation	87 (19)	2 (<1)	0	45 (19)	0	0
Abdominal pain	83 (18)	7 (1)	1 (<1)	60 (25)	10 (4)	0
Weight loss	81 (17)	5 (1)	0	14 (6)	0	0
Increase in ALT level	80 (17)	23 (5)	0	13 (5)	5 (2)	0
Mucosal inflammation	65 (14)	8 (2)	0	5 (2)	1 (<1)	0
Pyrexia	64 (14)	0	0	24 (10)	1 (<1)	0
Upper abdominal pain	63 (13)	3 (1)	0	31 (13)	0	0
Cough	63 (13)	1 (<1)	0	26 (11)	0	0
Peripheral oedema	63 (13)	4 (1)	0	32 (14)	2 (1)	0
Stomatitis	63 (13)	8 (2)	0	5 (2)	0	0
Dyspnoea	58 (12)	15 (3)	0	24 (10)	1 (<1)	0
Rash	58 (12)	2 (<1)	0	14 (6)	1 (<1)	0
Ascites	57 (12)	17 (4)	1 (<1)	30 (13)	11 (5)	0
Dysgeusia	56 (12)	0	0	5 (2)	0	0
Hypoalbuminemia	55 (12)	2 (<1)	0	12 (5)	0	0
Headache	52 (11)	1 (<1)	0	16 (7)	1 (<1)	0
Headache	52 (11)	1 (<1)	0	16 (7)	1 (<1)	0
Insomnia	49 (10)	1 (<1)	0	17 (7)	0	0
Dizziness	48 (10)	2 (<1)	0	15 (6)	0	0
Dyspepsia	47 (10)	0	0	7 (3)	0	0
Anaemia	46 (10)	18 (4)	1 (<1)	19 (8)	12 (5)	0
Back pain	46 (10)	5 (1)	0	24 (10)	1 (<1)	0
Increase in serum bilirubin level	45 (10)	10 (2)	4 (1)	17 (7)	2 (1)	2 (1)
Decrease in platelet count	45 (10)	13 (3)	4 (1)	7 (3)	2 (1)	0

***** Listed are adverse events, regardless of causality, that were reported in at least 10% of patients in either group. Severity was graded according to National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0.
Abbreviations: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; PPES, palmar-plantar erythrodysaesthesia syndrome.
Source: Abou-Alfa et al., 2018 (6).

The rate of discontinuation of cabozantinib or placebo owing to adverse events that were considered to be related to the trial regimen was 16% (76 patients) in the cabozantinib group and 3% (7 patients) in the placebo group. Adverse events leading to treatment discontinuation in more than 1.0% of patients in the cabozantinib group were palmar-plantar erythrodysesthesia, fatigue, decreased appetite, diarrhoea, and nausea (6).

AEs of any grade regardless of causality were reported in 99% of the patients in the cabozantinib group and in 92% in the placebo group, and AEs of grade 3 or 4 were reported in 68% of the patients in the cabozantinib group and in 36% in the placebo group (Table 34). The most common grade 3 or 4 AEs in the cabozantinib group were palmar-plantar erythrodysesthesia (17%, vs. 0% with placebo), hypertension (16% vs. 2%), increased aspartate aminotransferase level (12% vs. 7%), fatigue (10% vs. 4%), and diarrhoea (10% vs. 2%). The most common AEs of any grade leading to dose reductions in the cabozantinib group were palmar-plantar erythrodysesthesia (22%), diarrhoea (10%), fatigue (7%), hypertension (7%), and increased aspartate aminotransferase level (6%). Serious AEs were reported in 50% of the patients who received cabozantinib and in 37% of the patients who received placebo. A serious AE was defined as an AE of any grade that caused death, was life-threatening, resulted in hospitalisation or prolongation of hospitalisation, was deemed medically important, or resulted in disability or birth defect. Grade 5 AEs occurring within 30 days after the last dose of cabozantinib or placebo were reported in 55 patients (12%) in the cabozantinib group and in 28 patients (12%) in the placebo group and were commonly related to disease progression (6).

Grade 5 AEs that were considered to be related to cabozantinib or placebo were reported in 6 patients in the cabozantinib group (one event each of hepatic failure, tracheoesophageal fistula, portal-vein thrombosis, upper gastrointestinal haemorrhage, pulmonary embolism, and the hepatorenal syndrome) and in 1 patient in the placebo group (hepatic failure) (6).

B.2.10.2 Overview of the safety of the technology in relation to the decision problem

Cabozantinib has been licensed and marketed in the US since 2016, in Europe since 2016 for renal cell carcinoma and for HCC since November 2018. AEs in patients participating in the CELESTIAL trial were as expected in those with pre-treated advanced HCC. AEs characteristic of HCC in the context of chronic liver disease/cirrhosis were observed with cabozantinib and placebo and Grade 3 and 4 AEs associated with advanced HCC or underlying liver disease were reported frequently.

It is anticipated that cabozantinib will have an acceptable, recognisable, and manageable safety profile when used in the context of the decision problem.

Further details of AEs reported in the CELESTIAL study are provided in Appendix F.

B.2.11. Ongoing studies

No relevant studies are underway that are anticipated to provide additional evidence within the next 12 months or later to support the use of cabozantinib for the treatment of advanced HCC.

B.2.12. Innovation

Cabozantinib and regorafenib belong to the same drug class of TKIs. They inhibit multiple receptor tyrosine kinases (RTKs) implicated in tumour growth, metastasis, and angiogenesis, including VEGFR, angiopoietin receptor (TIE-2), mast/stem cell growth factor (KIT) and rearranged during transfection (RET). Cabozantinib is currently the only therapy developed for HCC that inhibits the mesenchymal epithelial transition factor (MET) and AXL receptors (in addition to VEGFR 1, 2 and 3), and thereby provides additional inhibitory effects beyond that of currently approved TKIs (8). Due to this unique molecular pathway, cabozantinib may be able to break TKI resistance established in the first line of treatment (41-43). Therefore, cabozantinib has a biologically plausible rationale to treat patients who are resistant to sorafenib. Thus, the proposed treatment pathway offers an additional treatment option for UK patients with advanced HCC, where systemic treatment options are limited and the prognosis remains poor as they continue to progress rapidly and have a short overall survival of 8 to 11 months (24, 40).

B.2.13. Interpretation of clinical effectiveness and safety evidence

Cabozantinib is indicated as monotherapy for the treatment of HCC in adults who have previously been treated with sorafenib. The proposed positioning of cabozantinib as a treatment option after prior treatment with sorafenib offers an alternative treatment option to a UK patient population with poor prognosis where there is only one other treatment option currently recommended by NICE. For these patients, cabozantinib offers an additional treatment option, including patients intolerant to sorafenib.

Cabozantinib is an oral multi-targeted inhibitor of RTKs that delivers significantly extended survival and delayed disease progression in patients with advanced HCC who have received prior therapy. This is supported by a robust, high quality phase 3 clinical programme as well as with indirect evidence versus regorafenib (the comparator in this submission) in the form of a Bucher ITC and MAIC.

The CELESTIAL trial was an international, randomised, double-blind, placebo-controlled, phase III trial. In the CELESTIAL trial, at the cut-off date for the second interim analysis of OS (01 June 2017), there was high maturity with a total of 484 deaths (78% actual information fraction) reported. The trial shows cabozantinib significantly reduced the risk of death by 24% compared with placebo and significantly reduced the risk of disease progression/death by 56% compared with placebo. Cabozantinib was associated with a significantly higher ORR than placebo. Consequently, cabozantinib was also associated with a lower rate of PD compared with placebo (21% versus 55%).

A key strength of the study was the inclusion of both second and third-line patients (28% of trial patients were receiving third-line therapy) and patients intolerant to sorafenib which is more reflective of real-world clinical practice and adds generalisability of the results to the UK population. This is in contrast to the RESORCE trial which provides clinical evidence for the regorafenib comparator. The RESORCE trial which only included patients who had received sorafenib first-line only i.e. the regorafenib population were pure second-line. Furthermore, the RESORCE trial included only patients who had disease progression on sorafenib and had to have tolerated sorafenib (≥400 mg daily for at least 20 of the 28 days before discontinuation);

The benefits of cabozantinib were accompanied by a manageable safety profile, as illustrated by patients in the cabozantinib group staying on treatment for almost twice as long as those in the placebo group (3.8 versus 2.0 months). Many AEs were as expected in patients with pre-treated advanced HCC, reflected by their high frequency in both the placebo and cabozantinib groups. The most frequently reported AEs in the cabozantinib group were typical of those with VEGFR-TKI therapies such as regorafenib (40) and consistent with the known safety profile of cabozantinib in

patients with advanced renal cell carcinoma (10). This is further supported by RWE studies such as those discussed in Table 32.

Conclusions from the evidence of the cabozantinib phase 3 clinical trial programme are supplemented by indirect comparisons designed to compare cabozantinib to regorafenib which was not included in the trial programme, but is relevant to National Health Service (NHS) clinical practice. Across these analyses, cabozantinib demonstrated comparable efficacy and a similar safety profile to regorafenib. This was shown through the conflicting direction of treatment benefit of the point estimates for OS and PFS. The confidence intervals showed that this was not statistically significant for OS. However, for PFS certain analyses showed a statistically significant treatment benefit for cabozantinib over regorafenib. Time-varying hazard ratio analyses showed that there was no divergence in treatment effect between the treatments over time. Additionally, evidence from the ITCs confirmed the rates of AEs are comparable across treatments.

There are existing uncertainties in the ITC which have been explored through a range of modelling techniques designed to establish the comparative treatment effect between cabozantinib and regorafenib. There was evidence to suggest that all the heterogeneity between the trials could not be accounted for, thus a conservative assumption of non-superiority is assumed, especially for the PFS endpoint as this favours cabozantinib. This assumption is in line with clinical expert feedback received during an advisory board (3) and responses received by NICE from professional bodies to the scoping consultation (5)

B.2.14. End of life criteria

Cabozantinib is not classified as a ‘life extending treatment at the end of life’ by NICE criteria.

B.3. Cost-effectiveness

B.3.1. Published cost-effectiveness studies

An SLR was conducted to obtain all published economic evaluation studies in the population under consideration (including studies reporting utility values and studies reporting cost and resource use data). Full details of the search are provided in Appendix G.

The SLR was originally conducted in April 2018 and an update search was performed in February 2021. The economic SLR identified a total of 71 studies described in 73 publications. Of the 71 economic evaluations, 62 studies were cost-effectiveness analysis (CEA), 5 studies reported cost-utility analysis (CUA) and 4 studies reported budget impact analysis. Cost-effectiveness, cost-utility studies, and budget impact studies are detailed in Appendix G.

Variation existed in modelling methodology across included publications with respect to study perspectives, sources of cost data, and approaches to modelling utilities. Modelled health states, source data for clinical inputs, and methods to extrapolate survival beyond the time horizon were generally similar across studies, with few exceptions. Across studies identified there was consistent use of a 3 health state model (progression-free, progressed and death) and a Markov or partitioned survival approach to calculating health state membership.

The majority of studies appropriately defined the advanced HCC study population and interventions. 49 studies clearly stated the perspective of the economic evaluation, reflecting the good applicability. However, costs and outcomes from other sectors were not appropriately measured and valued in all the economic evaluations.

A summary of modelling methodology across the relevant advanced HCC cost-effectiveness studies is presented in Table 35.

Table 35. 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)
Parikh 2017 (70)	2017	This cost effectiveness analysis used a Markov model consisting of 3 health states (PFS, PD, and Death) with a cycle length of 1 week at a 3% discount rate. Effectiveness data was obtained from Published clinical trial data and literature review. Cost data was obtained from Red book.	Unresectable HCC and Child-Pugh A cirrhosis and ECOD PS 0, 1	Regorafenib: 0.81 BSC: 0.63	Regorafenib: $47,112 BSC: $7408	$224,362/ QALY
Soto-Perez-de-Celis 2019 (71)	2019	This cost effectiveness analysis used a decision-analytic model. Discount rates of 10%, 20%, and 30% on the price of cabozantinib were included in Deterministic Sensitivity Analyses. Patients were classified into 3 mutually exclusive health states (progression-free disease, post progression disease and death). Effectiveness data was obtained from the area under the curve of progression-free survival and OS outcomes reported in the CELESTIAL RCT for both cabozantinib and placebo. The cost of each day of therapy with cabozantinib was determined from the 2018 Medicare Part D maximum allowed cost obtained using the previously published Memorial Sloan Kettering Cancer Centre	Incurable HCC, Child-Pugh class A liver function, progressive disease after treatment with sorafenib, and ECOG performance status of 0 or 1)	Cabozantinib: 0.75 Placebo + BSC: 0.68	*Cost in USD* **Drug costs** - Cabozantinib: $64,599 - Placebo + BSC: $0 **AEs cost** - Cabozantinib: $1,137 - Placebo + BSC: $207 **Cost of Post-progression therapies** - Cabozantinib: $35,290 - Placebo + BSC: $30,702 **Cost of EoL care** - Cabozantinib: $5,185	Cabozantinib vs Placebo + BSC: $1,040,675/QALY

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		Drug Abacus methodology. Costs of AEs were calculated according to published data from patients receiving treatment of various neoplasms. The costs of post-progression therapy were calculated according to the number of patients receiving each post-progression drug and/or intervention listed in the CELESTIAL trial. The duration of each post-progression therapy was obtained from published phase II/III trials, and the cost was obtained from the 2018 Medicare Part B or D maximum allowed cost depending on each drug. The cost of local therapy with embolisation was obtained from the 2018 Medicare Physician Fee Schedule			- Placebo + BSC: $5,448 **Monitoring cost** - Cabozantinib: $3,384 - Placebo + BSC: $3,384 **Total cost** - Cabozantinib: $109,596 - Placebo + BSC: $39,741	
Liao 2019 (72)	2019	This cost effectiveness analysis used a Markov model using TreeAge Pro 2011 (TreeAge Software) to simulate patients with sorafenib-resistant HCC receiving either cabozantinib or best supportive care. All costs and health outcomes were discount ed at 3% per year. Patients were classified into 3 mutually exclusive health states (progression-free disease,	Advanced sorafenib-resistant HCC	Cabozantinib: 0.61 Placebo + BSC: 0.48	*Cost in USD* **Incremental cost Cabozantinib vs Placebo + BSC** **• USA** - Full cost (Base case): $108,521 - 50% cost: $55,535 - 30 % cost: $34,340	Cabozantinib vs Placebo + BSC **• USA** - Full cost (Base case): $833,497/QALY - 50% cost: $426,532/QALY

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		progression disease and death). Monthly transition probabilities between health states were calibrated to best fit the Kaplan–Meier progression-free and overall survival curves from the CELESTIAL trial. EQ-5D index scores were obtained from literature. Cost of cabozantinib in USA, UK and China were obtained from AWP in the Red Book, published literature and Hong Kong list price respectively. Cost of Computed tomography imaging in USA and UK were obtained from published literature and West China Hospital in China Costs for managing grade 3-4 AEs in USA, UK and China were obtained from Red Book, British National Formulary and West China Hospital respectively. Costs for managing grade 3-4 AEs (PPE) in USA, UK and China were obtained from Local estimates			- 20% cost: $23,742 - 15% cost: $18,444 - 10% cost: $13,145 • **UK** - Full cost (Base case): $39,604 - 50% cost: $20,21 - 30 % cost: $12,188 - 20% cost: $8,272 - 15% cost: $6,314 - 10% cost: $4,355 • **China** - Full cost (Base case): $20,368 - 50% cost: $10,383 - 30 % cost: $6,389 - 20% cost: $4,392 - 15% cost: $3,393 - 10% cost: $2,395	- 30 % cost: $263,747/QALY - 20% cost: $182,354/QALY - 15% cost: $141,657/QALY - 10% cost: $100,961/QALY • **UK** - Full cost (Base case): $304,177/QALY - 50% cost: $153,775/QALY - 30 % cost: $93,613/QALY - 20% cost: $63,533/QALY - 15% cost: $48,493/QALY - 10% cost: $33,452/QALY

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
						• **China** - Full cost (Base case): $156,437/QALY - 50% cost: $79,747/QALY - 30 % cost: $4,970/QALY - 20% cost: $33,732/QALY - 15% cost: $2,663/QALY - 10% cost: $18,394/QALY
Shlomai 2019 (73)	2019	This cost effectiveness analysis used a Markov model using TreeAge Pro 2018 software and statistical analyses were performed in MATLAB. Annual discounting of all costs was done at a rate of 3%. The model consists of three health states (patients on Cabozantinib, Best supportive care or Death). Utility and Disutilities of Adverse Events were obtained from published	Advanced HCC who had failed prior treatments	Cabozantinib 60 mg daily: 0.86 Cabozantinib 36 mg daily: 0.86 Placebo: 0.70	*Cost in USD* - Cabozantinib 60 mg daily: $76,407 - Cabozantinib 36 mg daily: $47,614 - Placebo: $1 **Incremental cost of Cabozantinib 60 mg daily vs Placebo:** $76,406 **Incremental cost**	Cabozantinib 60 mg daily vs Placebo: $469,375/QALY Cabozantinib 36 mg daily vs Placebo: $292,496/QALY

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		literature. Cost of cabozantinib per 28-day cycle and treatment of the relevant AEs were based on drug prices taken from GoodRX on 21 October 2018			**of Cabozantinib 36 mg daily vs Placebo:** $47,613	
Shlomai 2018 (74)	2018	This cost effectiveness analysis used a Markov model in TreeAge Pro 2018 software and statistical analyses were performed in MATLAB. Annual discounting of the costs and benefit in this analysis was at a rate of 3%. The four health states considered were: Progression, Death, Patients live with AE and Patients live without AE. The overall mortality rate and Health states utilities were derived from RESORCE trial. Disutilities associated with AEs were derived from literature. Unit price of regorafenib was obtained from 2017 prices from GoodRX. AE costs were taken from Medicare physician fee schedule for 2017. Outpatient physician visits fees were obtained from current procedural terminology codes	Patients with advanced HCC and Child-Pugh A cirrhosis who had progressed on sorafenib	Placebo: 0.63 Regorafenib 120mg: 0.88 Regorafenib 144mg: 0.88 Regorafenib 160mg: 0.88	*Cost in USD* **Total Incremental Cost per patient** - Regorafenib (120 mg) vs Placebo: $50,022 - Regorafenib (144 mg) vs Placebo: $60,003 - Regorafenib (160 mg) vs Placebo: $66,558 **Incremental monthly cost:** Regorafenib (120 mg): $11,410 Regorafenib (160 mg): $15,186	Regorafenib (120 mg) vs BSC: $201,797/QALY Regorafenib (144 mg) vs BSC: $242,063/QALY Regorafenib (160 mg) vs BSC: $268,506/QALY
Upadhyay 2019 (75)	2019	This cost effectiveness analysis used a partitioned survival model with three health states (stable/progressed/death). Clinical inputs were obtained	Advanced HCC	Pembrolizumab vs Regorafenib: 0.08 Pembrolizumab vs Cabozantinib: 0.03	*Cost in USD* **Incremental costs:** - Regorafenib vs Pembrolizumab:	NR

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		from the KEYNOTE-224, RESORCE and CELESTIAL trials conducted for pembrolizumab, regorafenib and cabozantinib respectively. Cost, health state utility and AEs' disutility were obtained from public databases and published literature			$6,313 - Cabozantinib vs Pembrolizumab: $7,462	
Kim 2018 (76)	2018	This cost effectiveness analysis used a Markov model simulated using the clinical data of RESORCE. Possible health transitions reflected three states (Stable disease, Progressive disease, and Death). A 3%-time discount rate was used in the societal perspective analysis and 7% in the third-party payer's analysis	Advanced HCC	Regorafenib: 0.51 BSC: 0.39	*Cost in USD* **Total cost (societal perspective):** - Regorafenib: $65,901 - BSC: $32,467	Regorafenib vs BSC (societal perspective): $277,463/QALY
Sieg 2020 (77)	2020	This cost effectiveness analysis used a Markov model implemented in TreeAge Healthcare Pro 2019 software. The discounting of costs and utilities was performed with a rate of 3%. Model consists of three health states (stable, progressive and dead). Clinical data were obtained from published material of the CELESTIAL trial and the submitted GBA dossier of IPSEN Pharma and completed by a literature review on	Target population in the model was based on the CELESTIAL trial subjects. Adult patients with HCC who showed progression under prior sorafenib therapy, with Child-Pugh A liver function	Cabozantinib: 0.15	*Costs in USD* • **Germany** **Drug acquisition cost** - Cabozantinib: $53,018 - BSC: $0 **Adverse events** - Cabozantinib: $1,607 - BSC: $375 **Consultation** - Cabozantinib:	Cabozantinib vs BSC: **German model** -$306,778/LY -$375,470/QALY **United States model**: -$972, 049/LY - $1,189,706/QALY

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		cabozantinib, TKIs and HCC. In Germany model, DRG values were estimated using the DRG-Research Group Webgrouper. Drug prices and reimbursement amounts were deduced from the pharmacy database Lauer-Taxe of 15th April 2019. Study incorporated the current AMNOG amount of cabozantinib. In United states model, author determined the model costs using the US drug price portal GoodRX.com via extracting the average cash prices in April 2019. Study estimated physician outpatient fees, other services and hospitalisations using the 2019 physician fee schedule, clinical laboratory fee schedule and Medicare-Severity DRG classifications and software (HCPCS-DRG V1.0 Software) of Centres for Medicare and Medicaid Services and the methods of a published study			\$513 - BSC: \$434 **Laboratory** - Cabozantinib: \$202 - BSC: \$173 **Imaging** - Cabozantinib: \$1,281 - BSC: \$1,083 **Total** - Cabozantinib: \$56,621 - BSC: \$2,064 **Incremental Costs of Cabozantinib vs BSC** - Cabozantinib: \$53,018 - Adverse events: \$1,232 - Consultation: \$69 - Laboratory: \$29 - Imaging: \$198 - Total: \$54,556 • **United states** Drug acquisition cost - Cabozantinib:	

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
					$167,288 - BSC: $0 **Adverse events** - Cabozantinib: $6030 - BSC: $1075 **Consultation** - Cabozantinib: $1075 - BSC: $914 **Laboratory** - Cabozantinib: $868 - BSC: $751 **Imaging** - Cabozantinib: $2236 - BSC: $1890 **Total** - Cabozantinib: $177,496 - BSC: $4630 **Incremental Cost of Cabozantinib vs BSC** - Cabozantinib: $167,288 - Adverse events: $4,955	

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
					- Consultation: $161 - Laboratory: $117 - Imaging: $346 - Total: $172,866	

Abbreviations: QALYs, quality-adjusted life years; ICER, incremental cost-effectiveness ratio

B.3.2. Economic analysis

None of the published economic models compared the cost-effectiveness of cabozantinib versus regorafenib in adults who have previously been treated with sorafenib. A similar structure to the previous model submitted to NICE was adopted for the current submission as this standard 3 health state partitioned survival model is well-established in oncology modelling. Published cost-effectiveness studies identified in section B.3.1 use a similar 3 state partitioned survival model. Furthermore, this structure has been considered appropriate by NICE in advanced HCC. Therefore, a de novo model was developed using Microsoft Excel® (Office 365, version 2108) with Visual Basic for Applications functionality to assess the cost-effectiveness of cabozantinib versus regorafenib. This cost-effectiveness model (CEM) was created in addition to a simple economic model that was used to conduct a cost-comparison analysis (CCA). Details of the CCA are described previously in the fast track appraisal (FTA) document B.

B.3.2.1 Patient population

The de novo analysis assesses cabozantinib in adult patients with advanced HCC who have received prior sorafenib treatment and progressed following at least 1 prior systemic treatment, in comparison to regorafenib. This population is consistent with the ITT population of study CELESTIAL, the NICE final scope for this appraisal, the decision problem and the marketing authorisation for cabozantinib.

B.3.2.2 Model structure

A partitioned survival model was developed for the CEM; this approach allows direct modelling of overall survival. The direct correspondence between time-to-event endpoints (OS, PFS and TTD) and the survival functions in the model determines state membership. This approach also allows utilisation of individual patient level data from the CELESTIAL study and output from the ITC. Similar modelling approaches were accepted by NICE in the previous appraisal of regorafenib in advanced HCC.

The partitioned survival model includes three mutually exclusive health states: progression-free, progressed disease and death (Figure 31). State membership is determined by a series of independently modelled non-mutually exclusive time-to-

event curves derived from the CELESTIAL study. The model utilises the area under, and the difference between, time-to-event curves to estimate patient distribution between the disease states of interest:

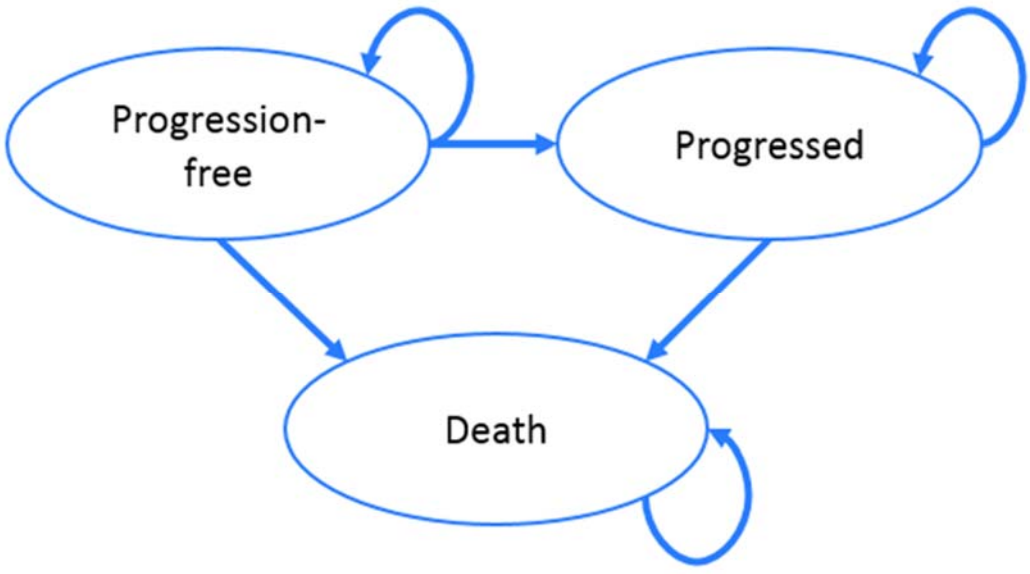
- Progression-free: All patients start treatment in the progression-free state. The proportion of patients who remain in the progression-free state was defined by PFS.

- Progressed disease: The proportion of patients with progressed disease was derived based on OS less PFS. Disease status was determined by the investigator using RECIST 1.1.

- Death: Death is an absorbing health state that patients enter from the progression-free and progressed disease states. The proportion of patients in the death state was derived as 1 less OS.

Each health state is associated with costs and utilities during the pre-defined time horizon.

Figure 31: 3 health state model structure diagram



The base case time horizon of the model is a lifetime (15 years), as recommended by NICE for treatments with a survival benefit (78). In practice, a time horizon with more

than 99% of patients who are modelled as having died will be considered an acceptable approximation of lifelong.

Costs and health outcomes are discounted at 3.5 % and the perspective of the NHS and personal social services is assumed, as per the NICE reference case.

The model cycle length is 28 days to reflect the dosing frequency of cabozantinib and regorafenib. This is consistent with the regorafenib NICE submission, where a 28-day cycle was modelled (1). This cycle length is considered short enough to represent the frequency of key clinical events. A half-cycle correction for outcomes is applied to reduce the potential for bias in the cost-effectiveness estimates.

Table 36 provides a summary of the features of the economic analysis as compared with previous appraisals in the population of interest.

Table 36. Features of the economic analysis

	Previous appraisals	Current appraisal	
Factor	**TA514 - regorafenib**	**Base case**	**Justification**
Time horizon	15 years	Lifetime (15 years)	In this disease, 15-years is effectively a lifetime time horizon which is appropriate in areas advanced HCC where differences in survival are expected.
Were health effects measured in QALYs; if not, what was used?	Yes	Yes	NICE reference case Only direct health effects related to patients were considered, and no wider societal impact or impact on carers
Discount of 3.5% for utilities and costs	Yes	Yes	NICE reference case
Perspective (NHS/PSS)	Yes	Yes	NICE reference case
Treatment waning effect?	None	None	Including survival benefits but excluding costs of treatment was not considered appropriate by the committee in TA555. This is explored in the model comparing the PFS and TTD endpoints and the effect on cost
Source of utilities	Based on EQ-5D data collected during the RESORCE study	Based on EQ-5D data collected in CELESTIAL study	EQ-5D-5L data were collected during the CELESTIAL study. It is the most appropriate data to use given it estimated utility values directly for patients considered within the submission. The EQ-5D-5L data were mapped to 3L using the Von Hout et al, as recommended by NICE (79). In addition, no other published values were found for a population with advanced HCC according to progression status.

	Previous appraisals	Current appraisal	
Factor	**TA514 - regorafenib**	**Base case**	**Justification**
			Scenario analyses using values from the regorafenib appraisal has been conducted.
Source of costs	The resource units as submitted by Bayer for the CDF reappraisal of sorafenib are used with updated unit costs	Physician survey based on 30 UK physicians treating advanced HCC patients NHS reference costs; PSSRU; BNF	As there is no real-world clinical experience relating to the use of cabozantinib in practice, we have conducted a survey of 30 UK practicing physicians, all with experience of treating more than 10 patients. In both the sorafenib and regorafenib appraisals there has been insufficient number of physicians' survey to elicit robust resource use estimates for advanced HCC patients in the UK.

Abbreviations: BNF, British National Formulary; CDF, Cancer Drug Fund; EQ-5D, EuroQol five-dimension; HCC: hepatocellular carcinoma; NHS, National Health Service; NICE, National Institute of Health and Care Excellence; PFS, progression-free survival; PSSRU, Personal Social Services Research Unit; QALY, quality-adjusted life-year; TTD, time to treatment discontinuation; UK, United Kingdom

B.3.2.3 Intervention technology and comparators

The final scope intervention is cabozantinib. As introduced in section B.1, regorafenib was selected as the only appropriate comparator because:

It is recommended by NICE for its licensed indication, adults with advanced unresectable HCC who have previously been treated with sorafenib. Cabozantinib has the same licensed indication and Ipsen are seeking the same positioning as the NICE recommendation for regorafenib. Since regorafenib is the only approved subsequent therapy for use after sorafenib, it is assumed to have a majority market share in this indication. This is supported by clinical experts estimation of regorafenib market share within the indication (3).

In post sorafenib patients eligible for treatment in the second and third-line setting, regorafenib is used in clinical practice. This is following the approval of atezolizumab plus bevacizumab in first-line, where sorafenib is now positioned as a second-line treatment option in addition to its use in first-line (4). Similarly, to second-line use, patients eligible for regorafenib in third-line are restricted to patients with ECOG performance status of 0 or 1 for which BSC is not a relevant treatment option. Therefore, regorafenib is the only comparator in this setting as patients would not be fit enough to receive chemotherapy. Consequently, regorafenib is the only appropriate comparator used in clinical practice which should form the basis for decision making.

B.3.3. Clinical parameters and variables

B.3.3.1 Incorporation of clinical data into the economic model

The primary source for clinical data in the economic model for the intervention is the Phase III pivotal randomised controlled trial, CELESTIAL, comparing cabozantinib to BSC. As regorafenib was not included in CELESTIAL, and there were no head-to-head trials comparing it to cabozantinib, an ITC was conducted to estimate its relative effectiveness (Section in B.2.9). The evidence from the ITC suggests equal efficacy between cabozantinib and regorafenib and so the base case is the CCA between cabozantinib and regorafenib. The equal efficacy assumption assumes that the only difference in treatment is the drug acquisition cost as the OS and PFS between treatments are equal. However, to assess the uncertainty in this assumption, a cost-

effectiveness sensitivity analysis was conducted using the CEM where the survival endpoints were modelled for the population adjusted anchored and unanchored comparisons included in the ITC.

Table 37 describes the survival data source and approach used for each analysis. Regorafenib OS, PFS and TTD were sourced as described in the ITC (Section B.2.9.2), from digitised KM curves from RESORCE and pseudo IPD generated, using the Guyot algorithm (53). In all scenarios the PFS, OS and TTD were extrapolated to the 15-year time-horizon of the model, as lifetime results are not available for patients in both studies (median follow-up of 22.9 months in CELESTIAL and 7.0 months in RESORCE). Guidance from the NICE DSU was followed to identify base case parametric survival models for OS, PFS and TTD (80). All parametric models were assessed against the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for statistical fit to the observed data. Curves were visually inspected and validated against relevant long-term data sources available to help identify the most plausible survival model. External long-term survival data in advanced HCC and clinical opinions were sought to validate the best fit models (3, 57).

Table 37. Summary of survival analyses

Analysis	Survival data	Modelling approach
Base case		
CCA – equal efficacy	Cabozantinib: CELESTIAL ITT cabozantinib arm Regorafenib: CELESTIAL ITT cabozantinib arm	Independent parametric models were fit to the OS and PFS curves from the CELESTIAL ITT cabozantinib arm. Regorafenib efficacy was assumed equal to cabozantinib. Details of curve fitting to the ITT CELESTIAL population are described in Appendix M and previously in the FTA document B.
Sensitivity analysis		
CEM – anchored MAIC, constant HRs	Cabozantinib: weighted CELESTIAL arms from MAIC (scenario 1) Regorafenib: both weighted CELESTIAL arms from MAIC (scenario 1), and RESORCE arms	Since the ITC anchored MAIC with a constant HR output was a Cox PH model i.e. a relative measure of effect (Section B.2.9.3), a base survival curve had to be generated to model absolute estimates of survival. It is theoretically incorrect to apply a HR derived from a different parametric model or from a Cox PH model to a base survival curve as per NICE guidance (80). Consequently, dependent PH models were used to apply a constant HR as follows: 1) Fit a parametric model to the weighted CELESTIAL data with treatment group as a covariate

Analysis	Survival data	Modelling approach
		2) Fit a parametric model to the RESORCE data with treatment group as a covariate 3) Apply the HR derived from Step 2 (the relative effect of regorafenib vs. placebo) to the weighted placebo arm of CELESTIAL to derive a placebo-adjusted survival curve for regorafenib
CEM – anchored MAIC, time-varying HRs	Cabozantinib: weighted CELESTIAL cabozantinib arm from MAIC (scenario 1) Regorafenib: both weighted CELESTIAL arms from MAIC (scenario 1) and RESORCE arms	As per the ITC anchored MAIC with time-varying HRs in Section B.2.9.3, independent parametric models were fitted to the weighted cabozantinib data, weighted CELESTIAL placebo data, regorafenib data and RESORCE placebo data in order to generate the hazard for each treatment arm. The time-varying cabozantinib versus CELESTIAL placebo HR was generated by dividing the hazard of the cabozantinib parametric model by the hazard of the CELESTIAL placebo parametric model at each timepoint. The regorafenib versus RESORCE placebo time-varying HR was generated in the same way. The time-varying HR of cabozantinib versus regorafenib was generated by calculating the ratio of the cabozantinib versus CELESTIAL placebo HR with the regorafenib versus RESORCE placebo time-varying HR. The cabozantinib arm was modelled using the independent parametric model fitted to the weighted cabozantinib data and the time-varying HR was applied to the survival curve for cabozantinib to generate the regorafenib curve.
CEM – unanchored MAIC	Cabozantinib: weighted CELESTIAL cabozantinib arm from MAIC (scenario 1) Regorafenib: RESORCE regorafenib arm	As per the results of the ITC unanchored MAIC in Section B.2.9.3, independent parametric models were fit to the OS and PFS curves from the weighted CELESTIAL cabozantinib arm from the MAIC and the RESORCE regorafenib arm.

Abbreviations: CCA, cost-comparison analysis; CEM, cost effectiveness model; FTA, fast track appraisal; HR, hazard ratio; ITC, indirect treatment comparison; ITT, intention-to-treat; MAIC, matching-adjusted indirect comparison; OS, overall survival; PFS, progression-free survival

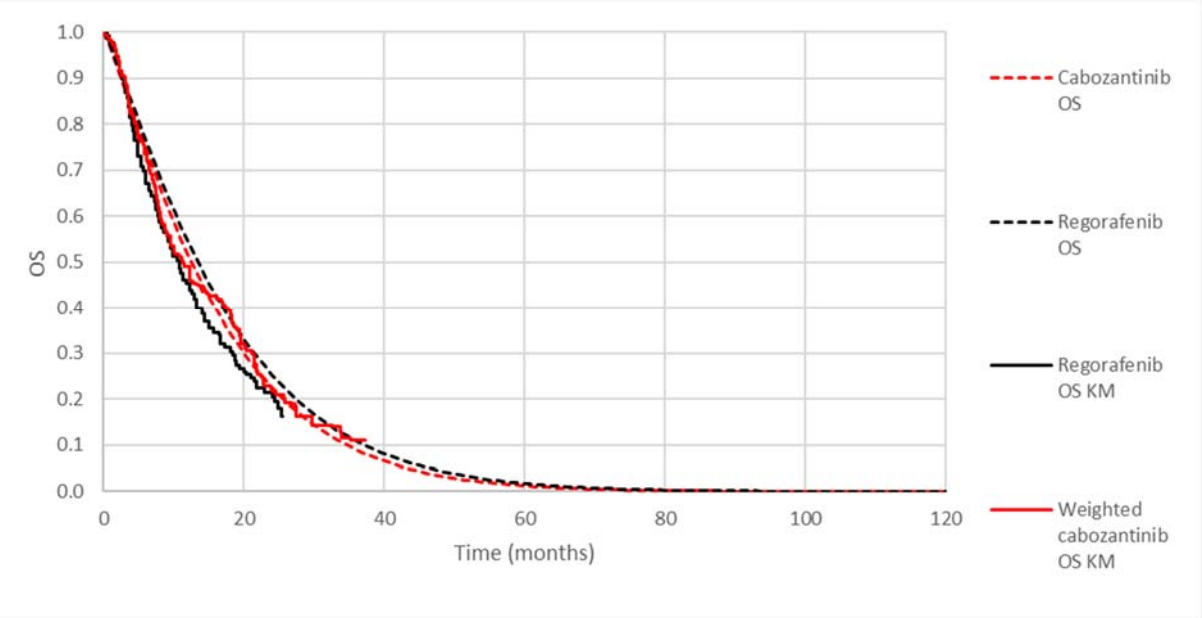
B.3.3.2 Overall survival

Anchored MAIC, constant HR scenario

The Weibull distribution was selected to model survival from the PH compatible parametric models. This model had the best statistical fit to the weighted CELESTIAL and RESORCE data (Table 38). The HR generated from a Weibull model for cabozantinib vs. placebo is 0.73 (95% CI: 0.52, 1.02) and for regorafenib vs. placebo is 0.67 (95% CI: 0.54, 0.83). The cabozantinib vs. regorafenib HR is 1.09 (95% CI:

0.73, 1.62) however, the examination of proportional hazards assumption in Section B.2.9.3 shows that the use of constant HR may not be appropriate for modelling OS. This is illustrated by the modelled regorafenib OS, which generates greater estimates than the regorafenib KM observed in the RESORCE trial, biasing the comparison against cabozantinib. The OS for cabozantinib and regorafenib is shown in Figure 32.

Figure 32: Cabozantinib and regorafenib OS generated from anchored MAIC constant HR (Weibull HR)



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; OS, overall survival

Table 38. AIC and BIC statistics for weighted CELESTIAL and RESORCE OS dependent parametric fits

Model	Weighted CELESTIAL		RESORCE	
	AIC	BIC	AIC	BIC
Exponential	2388.47	2396.84	2683.3	2692
Weibull	2376.85	2389.39	2661.89	2674.94
Gompertz	2389.54	2402.09	2680.73	2693.78

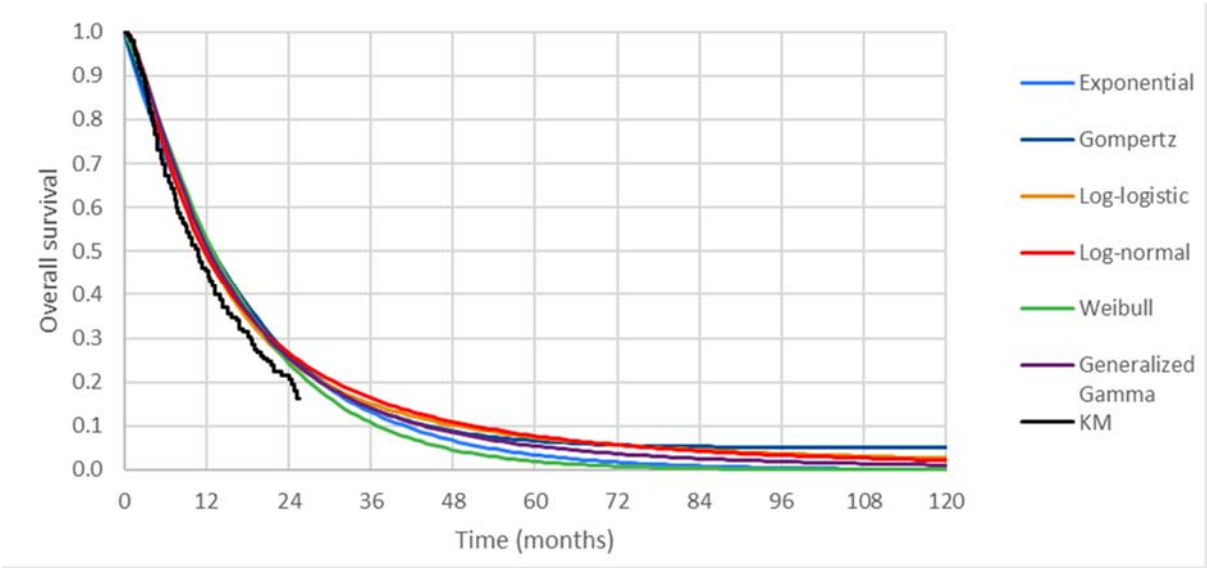
Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion.; ITT, intention-to-treat; OS, overall survival

Anchored MAIC, time-varying HR scenario

The log-logistic model was selected as the base case for this scenario as highlighted in Section B.2.9.3 (Anchored analysis using time-varying HR) and Appendix L, Section L.1.1. The time-varying HRs generated curve for regorafenib is shown in Figure 34

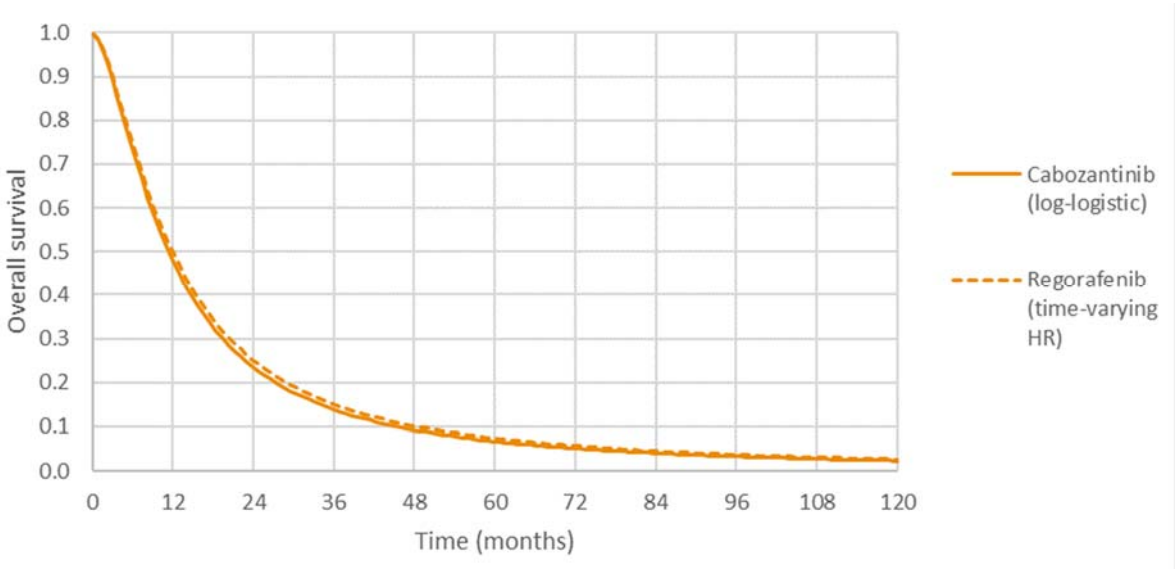
with the base survival curve for cabozantinib OS. Other base survival curves for cabozantinib OS and the resulting time-varying HR generated regorafenib OS curve can be seen in Figure 25 and Figure 33 respectively. The OS for regorafenib is closer to the observed values from RESORCE than the constant HR scenarios. However, the estimated OS is still greater than the OS KM from RESORCE after approximately 6 months

Figure 33: Regorafenib OS generated from anchored MAIC time-varying HR compared with RESORCE regorafenib OS KM



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; OS, overall survival

Figure 34: Cabozantinib and regorafenib OS from the anchored MAIC time-varying HR scenario



Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; OS, overall survival

Unanchored MAIC scenario

The log-logistic model was selected as the base case for this scenario as highlighted in Section B.2.9.3 (Unanchored analysis using independent parametric models). The parametric fits for cabozantinib and regorafenib OS are shown in Figure 25 and Figure 26.respectively. The log-logistic cabozantinib and regorafenib OS are shown in Figure 27.

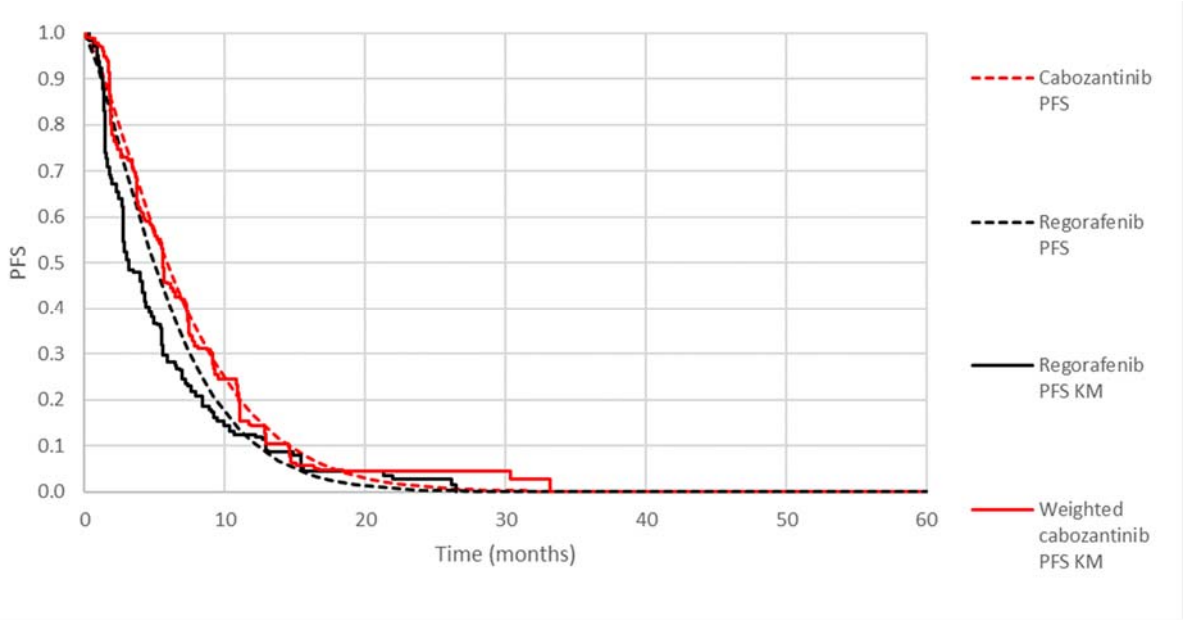
B.3.3.3 Progression-free survival

Anchored MAIC, constant HR scenario

The Weibull distribution was selected to model survival from the PH compatible parametric models. This model had the best statistical fit to the weighted CELESTIAL and RESORCE data (Table 39). The HR generated from a Weibull model for cabozantinib vs. placebo is 0.35 (95% CI: 0.26, 0.48) and for regorafenib vs. placebo is 0.44 (95% CI: 0.36, 0.53). The cabozantinib vs. regorafenib HR is 0.80 (95% CI: 0.55, 1.15). The examination of PH assumption in Section B.2.9.3 shows that the use of constant HR may not be appropriate for modelling PFS. Similarly to OS, this is illustrated by the modelled regorafenib PFS, which generates greater estimates than

the regorafenib KM observed in the RESORCE trial, biasing the comparison against cabozantinib. The PFS for cabozantinib and regorafenib is shown in Figure 35.

Figure 35: Cabozantinib and regorafenib PFS generated from anchored MAIC constant HR (Weibull HR)



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

Table 39. AIC and BIC statistics for weighted CELESTIAL and RESORCE PFS dependent parametric fits

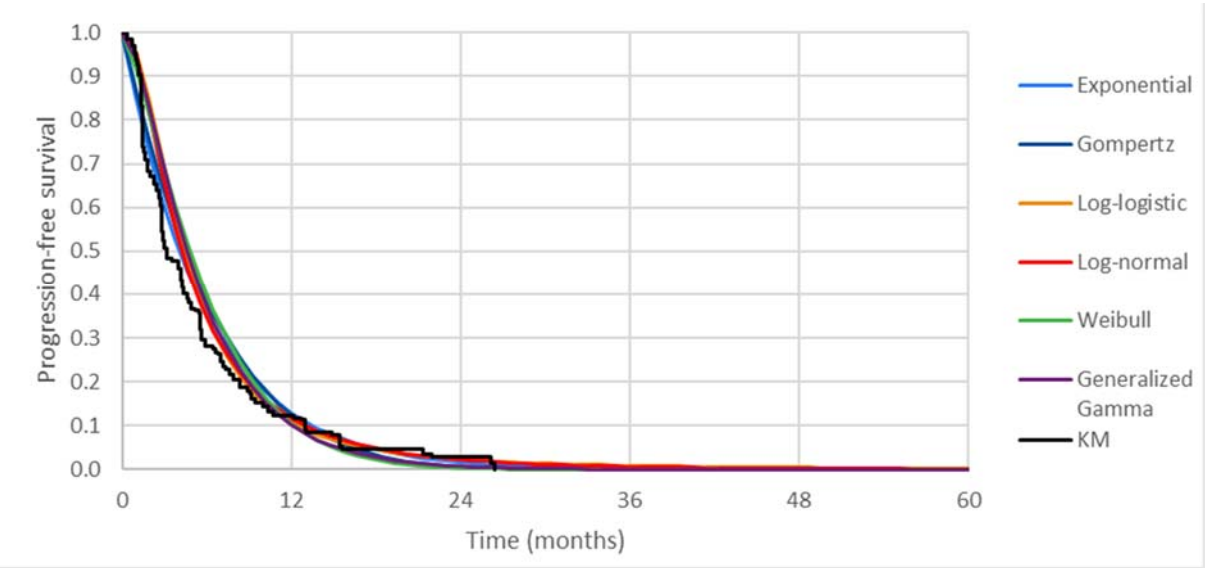
Model	Weighted CELESTIAL		RESORCE	
	AIC	BIC	AIC	BIC
Exponential	2026.14	2034.5	2373.66	2382.37
Weibull	1980.49	1993.03	2354.9	2367.95
Gompertz	2022.22	2034.77	2375.16	2388.21

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion.; ITT, intention-to-treat; PFS, progression-free survival

Anchored MAIC, time-varying HR scenario

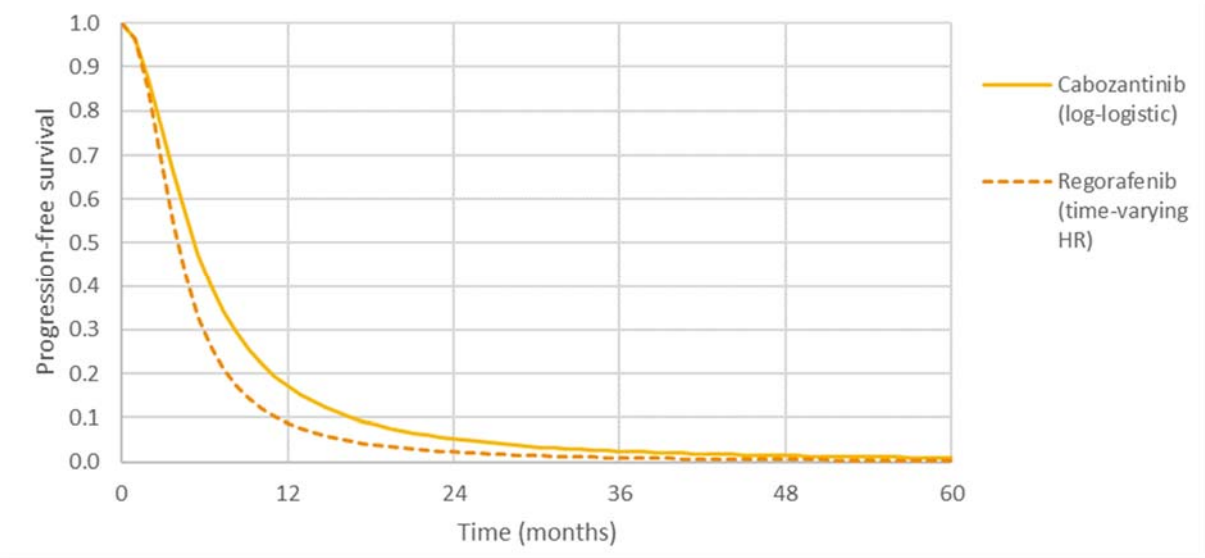
The log-logistic model was selected as the base case for this scenario as highlighted in Section B.2.9.3 (Anchored analysis using time-varying HR) and Appendix L, Section L.1.2. The time-varying HRs generated curve for regorafenib is shown in Figure 37 with the base survival curve for cabozantinib PFS. Other base survival curves for cabozantinib PFS and the resulting time-varying HR generated regorafenib PFS curve can be seen in Figure 28 and Figure 36 respectively.

Figure 36: Regorafenib PFS generated from anchored MAIC time-varying HR compared with RESORCE regorafenib PFS KM



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

Figure 37: Cabozantinib and regorafenib PFS from the anchored MAIC time-varying HR scenario



Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

Unanchored MAIC scenario

The generalised gamma model was selected as the base case for this scenario as highlighted in Section B.2.9.3 (Unanchored analysis using independent parametric

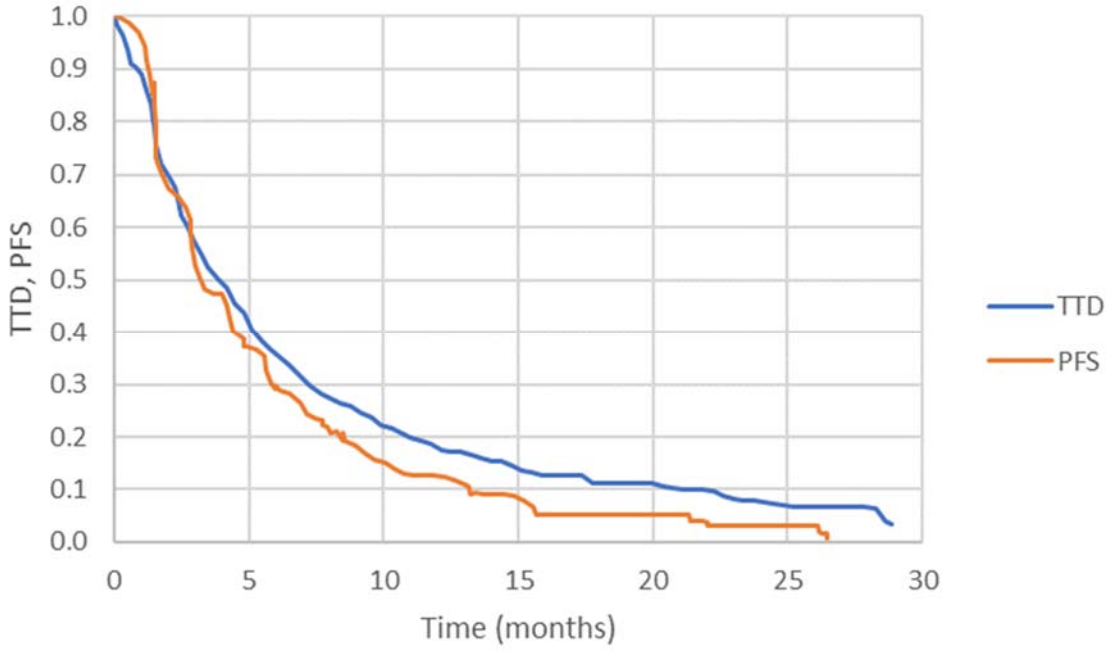
models). The parametric fits for cabozantinib and regorafenib PFS are shown in Figure 28 and Figure 29, respectively. The generalised gamma curves for cabozantinib and regorafenib PFS are shown in Figure 30Figure 30.

B.3.3.4 Time to treatment discontinuation

Both the CELESTIAL and RESORCE trials allowed treatment beyond progression; however, this was more pronounced for regorafenib as shown in the comparison of the TTD and PFS KM curves for both trials (Figure 38 and

Figure 39). It is possible this may have introduced a bias towards an improvement for regorafenib in terms of its OS endpoint as evaluated in the RESORCE trial. Patients are treated to progression and so cabozantinib and regorafenib TTD was modelled using PFS in the base case.

Figure 38: Comparison of the regorafenib treatment arm from RESORCE, PFS and TTD



Abbreviations: PFS, progression-free survival; TTD, time to treatment discontinuation
Note: patients on treatment on 29th February are considered censored
Source: Bruix et al. 2016; NICE TA555

Figure 39: Comparison of the cabozantinib treatment arm from CELESTIAL, PFS and TTD

Abbreviations: PFS, progression-free survival; TTD, time to treatment discontinuation
Source: CELESTIAL individual patient level data

A sensitivity analysis explored using the cabozantinib and regorafenib TTD curve from the 2L population of the CELESTIAL trial and the ITT population of the RESORCE trial in the MAIC adjusted scenarios. No population adjustment is applied to the TTD curves as there is uncertainty in the relationship between the clinical efficacy and TTD. The parametric fits to the cabozantinib TTD from the 2L population of CELESTIAL are shown in

Figure 40. The statistical fit is shown in Table 40. The long-term TTD extrapolations were validated by three clinical experts. From their experience in treatment patients with advanced HCC, these experts estimated that patients remaining on treatment at year 2, 3 and 4 will be 5%, 2% and 1% respectively (57). Comparison with external data and the statistical fit would suggest that the lognormal curve was the best fit for the TTD data.

Figure 40: Parametric fits for CELESTIAL 2L cabozantinib TTD

Abbreviations: 2L, second-line; KM, Kaplan-Meier; TTD, time to treatment discontinuation

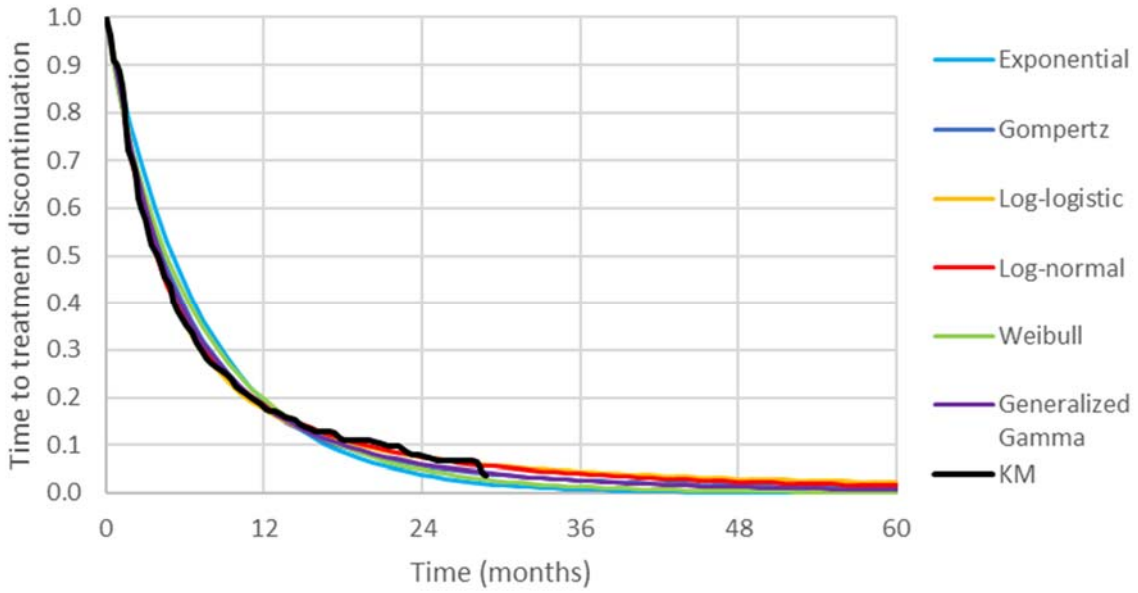
Table 40. AIC and BIC statistics for CELESTIAL 2L cabozantinib TTD parametric fits

Model	AIC	BIC
Exponential	893.64	897.44
Weibull	892.93	900.53
Gompertz	893.74	901.33
log-logistic	862.73	870.32
log-normal	852.78	860.37
Generalised gamma	853.00	864.38

Abbreviations: 2L, second-line; AIC, Akaike information criterion; BIC, Bayesian information criterion; TTD, time to treatment discontinuation

The parametric fits to the regorafenib TTD from the ITT population of RESORCE are shown in Figure 41 and the statistical fit is shown in Table 41. As per TA555, the log-logistic model was the best fitting model to the regorafenib TTD.

Figure 41: Parametric fits for RESORCE ITT regorafenib TTD



Abbreviations: ITT, intention-to-treat; TTD, time to treatment discontinuation

Table 41. AIC and BIC statistics for RESORCE ITT regorafenib TTD parametric fits

Model	AIC	BIC
Exponential	4703.30	4707.24
Weibull	4696.67	4704.54
Gompertz	4686.72	4694.60
log-logistic	4679.47	4687.35
log-normal	4684.62	4692.49
Generalised gamma	4678.20	4690.01

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion.; ITT, intention-to-treat; TTD, time to treatment discontinuation

B.3.3.5 Adverse reactions

The effect of including a different toxicity profile for cabozantinib and regorafenib was tested in the sensitivity analyses. The grade 3+ treatment-related AE incidences from the MAIC used in the model are shown in Table 42. These estimates for the incidence with cabozantinib had high uncertainty due to a low number of events available for analysis as discussed in Section B.2.10. This results in some AEs, such as hypertension, having a large point estimate.

The incidences were converted to a probability of an AE occurring per cycle while on treatment using the median time of exposure to treatment. This was [redacted] months for cabozantinib from CELESTIAL and 3.6 months for regorafenib from RESORCE (40, 81).This approach reflects the probability per cycle methodology used in TA555.

Table 42. AE grade 3 or more incidences included in sensitivity analysis

Adverse Events	Cabozantinib		Regorafenib	
	Incidence %	Probability per cycle (%)	Incidence %	Probability per cycle (%)
Palmar-plantar erythrodysaesthesia syndrome	13.2	3.0	12.6	3.7
Hypertension	55.2	15.8	13.1	3.8
Elevated aspartate aminotransferase	10.6	2.4	5.1	1.4
Fatigue	7.0	1.5	6.4	1.8
Diarrhoea	12.3	2.8	2.4	0.7
Elevated bilirubin	5.3	1.2	6.7	1.9

Abbreviations: AEs, adverse events
Source: Bruix et al,2017 (40).

B.3.4. Measurement and valuation of health effects

In the anchored and unanchored MAIC sensitivity analyses, a utility estimate for the 2L population was generated from the CELESTIAL 2L subgroup to assess the difference in QALYs between treatments.

B.3.4.1 Health-related quality-of-life data from the CELESTIAL study

EQ-5D-5L data was collected within the CELESTIAL study. Patients completed the EQ-5D-5L at baseline before any treatment, every 4 weeks until week 25, then every 8 weeks, irrespective of whether study treatment was given, reduced, interrupted, or discontinued, until the later of 8 weeks after investigator-determined radiographic disease progression per RECIST 1.1 or the decision to permanently discontinue study treatment. Patients did not receive medical results prior to completing the questionnaire. In contrast, the RESORCE trial collected EQ-5D-3L at day 1 of each treatment cycle. Patients were on treatment with regorafenib for the first 21 days of a 28 day cycle, therefore when the questionnaire was completed patients had been off treatment for 7 days. This may have biased health state utility and AE disutility

estimates from RESORCE for example the EQ-5D difference between cabozantinib and regorafenib is not reflective in the frequencies of grade 3/4 AEs. The utilities used in TA555 are tested in scenario analysis.

NICE recommends the use of the crosswalk approach to derive utility values for EQ-5D-5L health states in order to be aligned with previous valuations. Utility values used in crosswalk approach are derived from EQ-5D-3L valuation process.

NICE made a position statement that the EQ-5D-3L and the UK Time Trade Off (TTO) value set are the reference case for NICE submission. EQ-5D-5L data should be converted to EQ-5D-3L using the mapping function developed by van Hout et al. 2012 for the reference case analyses (79)

The EQ-5D-3L utility values for use in the cost-effectiveness model were mapped from EQ-5D-5L data collected from the CELESTIAL study.

B.3.4.2 Mapping

The EQ-5D-5L health states obtained by patients in different time points in CELESTIAL study were used to derive utility scores based on the EQ-5D-3L value sets for the UK. The utility values in the CELESTIAL study are based on value sets for the USA. Converting to UK value sets and 3L was performed using the 'crosswalk' developed by van Hout, et al. (79). This is the utility derivation method recommended by NICE for data gathered using the EQ-5D-5L (79). The crosswalk value sets used were developed by the EuroQol group.

As part of the sensitivity analysis, utility values were also derived using the algorithm based on EQ-5D-5L, not using the mapping or crosswalk to EQ-5D-3L utility values. This method is based on a scoring algorithm for the general population presented by the Office of Health Economics (Office of Health Economics, 2014). Preference-based valuation of EQ-5D-5L sets was conducted by OHE using a protocol developed by the EuroQol Group.

B.3.4.3 Health-related quality-of-life studies

A literature search (see Appendix H) was conducted to locate utility values that were suitable for inclusion in the economic model. The HRQoL SLR aimed to identify the

available HRQoL evidence for any interventions in the treatment of patients with advanced HCC. Values were required for pre-progression and progressed disease in a population of patients with advanced HCC. Several HRQOL publications reported quality of life values according to different instruments, but preference-based utility values were not reported, and these are not suitable for the economic model. Economic evaluations were available and the only source of utility values for patients with HCC was based on the phase III study of sorafenib vs lenvatinib (Hudgen 2018). The values from the sorafenib submissions to NICE and the SMC do report utility values according to the same health states for a comparable population of patients and, having been used before are the only other alternative values for use in the economic model. However, these utility values have a lack of face-validity as the progressed utility value is numerically higher than the pre-progressed utility value. For the purposes of the economic evaluation of cabozantinib the preferred values are those derived using the EQ-5D measure collected in the CELESTIAL study (see Health-related Quality of life data used in the cost-effectiveness analysis)

B.3.4.4 Adverse reactions

Based on the CELESTIAL study, grade 3/4 TEAEs were included in the model and only those TEAEs occurring in ≥5% of patients in either arm was included. This is a standard approach to including TEAEs, and was used in TA555, as grade 3/4 events are likely to be both costlier and have a greater impact on patient's quality of life than grade 1/2 events. The disutility in the model was calculated from the product of the probability per cycle of an AE occurring and the proportion on treatment multiplied by disutility.

B.3.4.5 Health-related quality-of-life data used in the cost-effectiveness analysis

EQ-5D-3L utility values were analysed using multiple model types, using both univariable and multivariable model structures, and clustering by time-point defined by in the cost-effectiveness model. The variables which were tested as independent variables in the models include:

- Treatment (still on treatment vs after treatment discontinuation/finalisation) at time of EQ-5D-5L completion

- Progression status (yes or no) at time of EQ-5D-5L completion
- TEAEs, defined as presence of grade 3-4 adverse events (yes vs no) at time of EQ-5D-5L completion

In addition to the inclusion of individual variables in the model, the interaction between some variables have been tested in order to assess their impact on the resultant utility values. For example, examining the interaction between treatment and TEAEs grade 3-4, establishes if utility values associated to the presence of TEAEs differs between patients still on treatment or patient after treatment discontinuation. Interaction terms which were not statistically significant were excluded from the model.

As noted above, to obtain EQ-5D-3L utility values from the CELESTIAL trial to be included in the cost-effectiveness model, different statistical models have been tested. The following types of regression models were tested:

- Ordinary Least Squares (OLS) regression - OLS model does not consider repeated EQ-5D-5L assessments for patients between study visits.
- Tobit regression with repeated measurements - The Tobit regression model has been previously used in other studies to derive utilities due to the presence of negative utility values (corresponding to health states worse than death). Using Tobit model negative utility values were transformed to 0.
- Mixed model for repeated measurements - Allows repeated EQ-5D-5L measurements at patient level to be considered given that patients provided several assessments during the study follow-up period.

The selection of the preferred model was defined based on the following criteria:

- Model reflecting the repeated nature of measurements
- Selection based on AIC measurements
- Smallest difference between the predicted and the observed values

Table 43 below presents the index scores generated from the various models.

Table 43. Index score EQ-5D-3L for the different tested models

		Univariate OLS	Multivariate OLS	Multivariate mixed model for repeat measurement	Multivariate Tobit mixed model for repeat measurement
Intercept			[redacted]	[redacted]	[redacted]
Treatment	Still on treatment	[redacted]	[redacted]	[redacted]	[redacted]
	Treatment discontinuation	[redacted]	[redacted]	[redacted]	[redacted]
Progression status	No	[redacted]	[redacted]	[redacted]	[redacted]
	Yes	[redacted]	[redacted]	[redacted]	[redacted]
TEAE grade >= 3	No	[redacted]	[redacted]	[redacted]	[redacted]
	Yes	[redacted]	[redacted]	[redacted]	[redacted]
AIC		[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike information criterion; TEAE, treatment-related adverse event

The multivariable OLS regression model has the lowest AIC (-4391.09). However, the model does not reflect the nature of data collected given that it does not consider repeated measures of EQ-5D health states between study visits. Due to the repeated measures at the patient level for different timepoints, this method is not considered the most appropriate model in this case. In addition, the number of questionnaires reported by each patient can be different and this can produce a bias in the results when using this model.

Multivariable Tobit model with repeated measures has been previously used to analyse utility variables in order to reflect the scale used for negative values, corresponding to health states worse than death, and the distribution can sometimes be left-skewed. Based on data obtained from the CELESTIAL study only approximately 1% of utility values correspond to negative values, having a lower impact on estimated utility values. Table 44 describes differences between predicted and observed utility values for EQ-5D-3L. Mixed model provided higher errors in the prediction of utility values, proving the Tobit model to have more accurate predictions.

Table 44. Difference between predicted and observed utility values for EQ-5D-3L for Tobit and mixed models for repeated measures

	n	Mean	Standard deviation	Minimum	Maximum
Tobit model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mixed model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Therefore, the multivariable Tobit regression with mixed model for repeated measurements model seems to be the best option, it has a lower AIC to those values obtained with the Mixed model (-1772.93 vs -1931.16), errors obtained with prediction are closer to zero. This model considers that each patient has a different number of questionnaires but does require imputation in response variable, by imputing all negative utility values into zero.

As mentioned above, once the most appropriate model was selected, it is then important to determine if any variables included in the regression analysis are confounding factors. Given that that both disease progression and treatment discontinuation are highly correlated, the selected multivariable Tobit regression with mixed model for repeated measurements was also obtained excluding treatment discontinuation.

Excluding treatment discontinuation, all independent variables included in the model were statistically significant, obtaining an AIC of -1935.72. All the variables were not statistically significant when both treatment discontinuation and progression status was included in the model, thus highlighting treatment discontinuation is a confounding factor and should not be included in the multivariable Tobit regression with mixed model for repeated measurements. Applying health state utilities represents the base case analysis as it is most representative of the way in which utility values have been incorporated in previous NICE submissions in advanced HCC (1).

The utility values used in the anchored and unanchored MAIC sensitivity analyses are presented in Table 45 below.

Table 45. Summary of utility values for anchored and unanchored MAIC sensitivity analyses

Health state	Utility value, mean	Standard error
Progression-free	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]
Disutility due to AE	[redacted]	[redacted]

Abbreviations: AE, adverse event

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

A SLR was conducted to identify healthcare resource use (HCRU) and accompanying cost associated with the proposed population. The SLR found 30 studies that reported costs associated with the treatment of advanced HCC, while resource use was reported in 17 studies. The search identified 2 publications on cost from the UK. Their brief overview is provided in Appendix I. A UK clinician survey was conducted and the survey was designed to elicit responses from 30 UK clinicians in order to accurately estimate resource use in current clinical practice in the UK (82).

B.3.5.1 Intervention and comparators' costs and resource use

The drug acquisition cost of cabozantinib is based on the PAS price per pack of [redacted]. The list price was used for regorafenib as the PAS price was unknown. The maximum daily dose of cabozantinib and regorafenib is 60mg and 160mg respectively, however, treatment could be interrupted, or the dose reduced, to help manage side effects. In the CELESTIAL trial, the mean daily dose of cabozantinib was 36.6 mg and in the RESORCE trial the mean daily dose of regorafenib was 144.1 mg (40, 44). The cost per model cycle accounted for the average dosing observed in the trials.

The use of relative dose intensity is in line with assumptions used in TA555 where the guidance indicates full pack dosing was "unlikely to reflect clinical practice, because the dose reductions in the trial were planned, so it was more likely that wastage would be minimised in clinical practice" (TA555 guidance, Section 3.15). In TA555, the NICE Appraisal Committee concluded that "although wastage could be minimised, the

pharmacists' evidence provided by the company suggested that it could not be eliminated entirely".

In the appraisal of sorafenib for previously untreated advanced HCC (TA474) it was considered that full pack dosing was not clinical practice, but that wastage could not be eliminated entirely (38). The cost of wastage was taken into account by implementing a one-off cost per patient that equalled a quarter of the cost of a course of treatment. This was taken from an exploratory analysis presented in TA474 that showed there was wastage of up to 7 days' worth of treatment of sorafenib (28 day cycle). A similar approach has been tested in scenario analysis.

Cabozantinib and regorafenib are given in combination with BSC, which includes various medications. These were estimated through the physician survey of 30 current clinicians in the UK (82). The drug acquisition costs are shown in Table 46 and the average cost per day used in the model in Table 47.

Table 46. Drug acquisition costs

Drug	Dose per day	Pack size	Pack price (£)	Reference
Cabozantinib (60 mg tablet)	60 mg RDI 61.0% =(36.6/60)	30	PAS: [redacted]	Ipsen
Regorafenib (40 mg tablet)	160 mg per day for 21/28 days RDI 90.1% =(144.1/160)	84	3,744	BNF (13)
Concomitant BSC				
Cyclizine hydrochloride (50 mg tablet)	150mg	100	3.40	Weighted average price eMIT 2021
Dexamethasone (4mg tablet)	8mg	50	12.99	
Lactulose (5ml soln)	30ml	500	1.84	
Metoclopramide (10 mg tablet)	30mg	28	0.35	
Morphine sulphate (1 mg/ml injection)	10ml	10	6.21	
Omeprazole (20mg tablet)	20mg	28	0.35	
Oramorph (10mg/5ml))	60mg	100	3.65	
Paracetamol (500 mg tablet)	4,000mg	32	0.22	

Drug	Dose per day	Pack size	Pack price (£)	Reference
Spironolactone (100mg tablet)	100mg	28	1.20	

Table 47. Average drug acquisition costs per day

Drug	Average cost per day
Cabozantinib	[redacted]
Regorafenib	£120.42
Concomitant BSC	£1.72

Abbreviations: BSC, best supportive care

B.3.5.2 Health-state unit costs and resource use

In both the sorafenib and regorafenib appraisals there was an insufficient number of physicians surveyed, to elicit robust resource use estimates for advanced HCC patients in the UK (1, 38). Therefore, the estimates on the resource use associated with the management of patients with advanced HCC were determined through another resource use survey which was conducted in June 2018 (82). The survey was based on 30 clinical experts in the field of oncology in the UK who have treated at least 10 advanced HCC patients in the last 12 months, all of whom were familiar with using sorafenib.

The health state resource use unit costs are presented in Table 48 and the total health state costs are shown in Table 49.

Table 48. Health state resource use unit costs

Variable	Unit Cost	Code, Details	Reference
Medical staff visits			
Oncologist	£204.48	Cost per visit	NHS National Schedule of Reference Costs 2019/20 (specialty code 370, weighted average WF01A-WF02C consultant led)
Hepatologist	£174.44	Cost per visit	NHS National Schedule of Reference Costs 2019/20 (specialty code 306, average WF01A-WF02B consultant led)
Gastroenterologist	£154.41	Cost per visit	NHS National Schedule of Reference Costs 2019/20 (specialty code 301, weighted average WF01A-WF02D consultant led)

Variable	Unit Cost	Code, Details	Reference
Clinical Nurse Specialist	£44.00	Cost per visit, assumed band 8b.	PSSRU, Unit Costs of Health and Social Care 2021. Nurse (GP practice). Cost per hour, including qualifications
Palliative Care Team	£44.00	Cost per visit	PSSRU, Unit Costs of Health and Social Care 2021. Nurse (GP practice). Cost per hour, including qualifications
Macmillian Nurse	£44.00	Cost per visit, assumed band 8b.	PSSRU, Unit Costs of Health and Social Care 2021. Nurse (GP practice). Cost per hour, including qualifications
GP	£39.00	Cost per 9.22-minute visit	PSSRU, Unit Costs of Health and Social Care 2021. General practitioner, cost per surgery consultation lasting 9.22 minutes (including direct care staff costs, with qualification costs)
Laboratory tests			
AFP test	£8.56	Cost per test	NHS reference costs 2019/20. weighted average of DAPS01 and DAPS02 (cytology, and histopathology and histology)
Liver function test	£8.56	Cost per test	NHS reference costs 2019/20. weighted average of DAPS01 and DAPS02 (cytology, and histopathology and histology)
Biochemistry	£1.20	Cost per test	NHS reference costs 2019/20. DAPS04 (clinical biochemistry)
Complete blood count	£2.27	Cost per test	NHS reference costs 2019/20. weighted average of DAPS03, DAPS05 and DAPS08 (integrated blood services, haematology and phlebotomy)
International normalized ratio (INR)	£2.27	Cost per test	NHS reference costs 2019/20. Average of DAPS03, DAPS05 and DAPS08 (integrated blood services, haematology and phlebotomy)
Radiological tests			
Computerised tomography (CT) scan (abdominal)	£123.71	Cost per test	NHS National Schedule of Reference Costs 2019-2020 (code RD22Z)
Magnetic resonance imaging (MRI) (abdominal)	£273.25	Cost per test	NHS National Schedule of Reference Costs 2019-2020 (code RD03Z)
Procedures			
Radiotherapy fraction	£739.30	Cost per procedure	NHS National Schedule of Reference Costs 2019/20 (code SC56Z)
Hospitalisations			
General ward	£676.48	Cost per day	NHS reference costs 2019/20, NHS reference costs 2015/16 non-

Variable	Unit Cost	Code, Details	Reference
			elective long-stay admissions, the mean cost per bed day weighted by the number of total finished consultant episodes
A&E admission	£205.09	Cost per admission	Average of codes: VB01Z-VB09Z and VB11Z. NHS reference costs 2019/20.
ICU	£270.61	Cost per day	NHS National Schedule of Reference Costs 2019/20 (code 315)

Abbreviations: A&E, accident & emergency; ICU, intensive care unit
Source: NHS National Schedule of Reference costs (83); PSSRU (84)

Table 49. Health state costs

Health state	Health state cost per cycle
Progression-free	
Hospitalisations	£624.02
Radiological tests	£204.10
Medical Staff Visits	£17.69
Lab tests	£71.76
Procedures	£8.92
Progressed disease	
Hospitalisations	£1,057.79
Radiological tests	£259.96
Medical Staff Visits	£14.76
Lab tests	£26.23
Procedures	£3.86
One-off cost at disease progression	
Lab tests	52.06
Radiological tests	575.81

B.3.5.3 Adverse reaction unit costs and resource use

The costs of AEs have been drawn from previous NICE appraisals such as the regorafenib appraisal (TA555); however recent discussions with two clinical experts (85) have demonstrated that they are now familiar with the AE profiles of TKIs such that most grade 3 AEs included can be managed via temporary cessation of treatment, dose reduction and supportive therapies. These AEs can often be managed via telephone discussion without the need for the patient to be seen in a hospital setting. The only grade 3 AE that clinical experts thought would warrant hospital admission would be grade 3 diarrhoea. Thus, the costs of managing AEs are in in reality likely to

be significantly lower. The grade 3+ treatment-related AE costs used in the model are shown in Table 50 and are those included in TA555 for consistency although in reality only grade 3+ AEs are likely to need hospitalisation as mentioned above. The impact of AE cost assumptions were tested in the sensitivity analysis. The AE cost was calculated as a one-off cost at the start of the model time horizon using the product of the incidence of the AE and unit cost.

Table 50: Adverse event costs

Adverse event	Cost per episode	Code, Details	Reference
Diarrhoea	£629.69	FD10K Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 6-10 - non-elective short stay	NHS reference costs 2019/20 (83)
Aspartate aminotransferase increase	£0.00	-	Based on the assumptions: regular blood tests (already considered under health state management costs)
Hypertension	£638.81	EB04Z Hypertension – Total HRG	NHS reference costs 2019/20 (83)
Fatigue	£63.45	Based on cost included in sorafenib NICE submission (38)	Inflated using PSSRU 2021 (84)
Palmar-plantar erythrodysaesthesia syndrome	£420.66	JD07J Skin Disorders without Interventions, with CC score 2-5 - non-elective short stay	NHS reference costs 2019/20 (83)
Elevated bilirubin	£0.00	-	Based on the assumptions: regular blood tests (already considered under health state management costs)

Abbreviations: NHS, National Health Service; PSSRU, Personal Socaial Services Research Unit

B.3.5.4 Miscellaneous unit costs and resource use

Cost associated with end-of-life terminal treatment was also included in the model. This was applied as one-off cost to those patients who died during the time horizon (Table 51).

Table 51. Terminal care costs

Variable	Unit Cost	Code, Details	Reference
Terminal care	£5,818.34	Average cost of hospital and hospice stays for patients with cancer.	Coyle et al. (86) Inflated using to PSSRU 2021 (84).

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

The base case analysis is the CCA where cabozantinib and regorafenib are assumed to have equal efficacy. The details of the CCA inputs and assumptions are described in the FTA. The inputs and assumptions used in the CEM sensitivity analyses are detailed in Table 52 and Table 53.

Table 52. Summary of key variables applied in the CEM

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Overall survival			
Anchored constant HR scenario	Cabozantinib parametric curve: Weibull Cabozantinib vs regorafenib HR: 1.09	Cabozantinib vs regorafenib HR: 0.73, 1.62 (multi-variate normal distribution for survival curves)	Section B.3.3.2
Anchored time-varying HR scenario	Cabozantinib parametric curve: Log-logistic Cabozantinib vs regorafenib HR: Log-logistic	(multi-variate normal distribution for survival curves)	Section B.3.3.2
Unanchored scenario	Cabozantinib parametric curve: Log-logistic Regorafenib parametric curve: Log-logistic	(multi-variate normal distribution for survival curves)	Section B.3.3.2
Progression-free survival			
Anchored constant HR scenario	Cabozantinib parametric curve: Weibull Cabozantinib vs regorafenib HR: 0.80	Cabozantinib vs regorafenib HR: 0.55, 1.15 (multi-variate normal distribution for survival curves)	Section B.3.3.2
Anchored time-varying HR scenario	Cabozantinib parametric curve: Log-logistic	(multi-variate normal distribution for survival curves)	Section B.3.3.2

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
	Cabozantinib vs regorafenib HR: Log-logistic		
Unanchored scenario	Cabozantinib parametric curve: Generalised gamma Regorafenib parametric curve: Generalised gamma	(multi-variate normal distribution for survival curves)	Section B.3.3.2
Utility in PF	[redacted]	[redacted] (Beta distribution)	Section B.3.4
Utility in PD	[redacted]	[redacted] (Beta distribution)	Section B.3.4
Disutility due to AE	[redacted]	[redacted] (Beta distribution)	Section B.3.4

Abbreviations: AE, adverse event; CEM, cost-effectiveness model; CI, confidence interval; HR, hazard ratio. PD, progressed disease; PF, progression-free

Table 53. Key assumptions used in the CEM

Area	Assumption/Setting	Justification
Time horizon	15 years	A lifetime horizon is appropriate for a condition where a survival difference is shown
Comparators included in the economic model	Regorafenib	Section B.1 & B.3.2.3
Treatment duration	TTD equal to PFS	Patients are treated to progression
RDI	Cabozantinib 61.0% Regorafenib 90.1%	The CELESTIAL and RESORCE trials efficacy results were obtained including dose reductions and treatment interruptions to manage adverse events. This treatment approach is in keeping with clinical practice where dose reductions/interruptions are a standard part of patient care and wastage minimised
Discount rate for costs and outcomes	3.5%	In line with NICE reference case
Drug acquisition costs	Cabozantinib: PAS price Regorafenib: list price	Given the confidentiality of PAS prices, a comparison was not feasible with the PAS price for regorafenib

Abbreviations: PAS, patient access scheme; PFS, progression-free survival; RDI, relative dose intensity; TTD, time to treatment discontinuation

B.3.7. Base case results

B.3.7.1 Equal efficacy base case

In the base case, cabozantinib and regorafenib have equal efficacy and only drug acquisition costs from the time on treatment are considered. This scenario is the revised base case from the FTA which uses preferred ERG assumptions for inclusion of RDI. The results of this base case are derived from the CCA and are shown in [redacted] to [redacted].

Table 54. Further details of this analysis are described in the FTA Document B and FTA ERG report. Including a quarter pack of wastage increased the cost savings from [redacted] to [redacted].

Table 54. Base case results

Technologies	Total costs
Cabozantinib	[redacted]
Regorafenib	£29,952
Difference	[redacted]

B.3.7.2 Anchored MAIC constant HR scenario

The results are shown in Table 55. The ICER is in the south west quadrant of the cost-effectiveness plane (less effective and lower costs) with a net monetary benefit (NMB) of £17,474 at £30,000 willingness to pay (WTP) per QALY. The incremental QALYs are very small and hence the ICER is unstable.

Table 55. Anchored MAIC constant HR scenario results

Technologies	Total costs	Total LYs	Total QALYs	ICER (£/QALY)
Cabozantinib	[redacted]	1.39	[redacted]	
Regorafenib	£55,669	1.48	1.04	
Incremental	[redacted]	-0.09	[redacted]	SW £290,383

Abbreviations: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; LY, life-year; MAIC, matching-adjusted indirect comparison; QALY, quality-adjusted life-year; SW, south west

B.3.7.3 Anchored MAIC time-varying HR scenario

The results are shown in Table 56. The ICER is in the south west quadrant of the cost-effectiveness plane (less effective and lower costs) with a NMB of £16,471 at £30,000 WTP per QALY. The incremental QALYs are very small and hence the ICER is unstable.

Table 56. Anchored MAIC time-varying HR scenario results

Technologies	Total costs	Total LYs	Total QALYs	ICER (£/QALY)
Cabozantinib	[redacted]	1.69	[redacted]	
Regorafenib	£60,496	1.78	1.25	
Incremental	[redacted]	-0.09	[redacted]	SW £300,170

Abbreviations: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; LY, life-year; MAIC, matching-adjusted indirect comparison; QALY, quality-adjusted life-year; SW, south west

B.3.7.4 Unanchored MAIC scenario

The results are shown in Table 57. The ICER is in the south east quadrant of the cost-effectiveness plane (more effective and lower costs) with a NMB of £17,837 at £30,000 WTP per QALY.

Table 57. Unanchored MAIC scenario results

Technologies	Total costs	Total LYs	Total QALYs	ICER (£/QALY)
Cabozantinib	[redacted]	1.69	[redacted]	
Regorafenib	£56,058	1.52	1.07	
Incremental	[redacted]	0.17	[redacted]	Dominant

Abbreviations: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; LY, life-year; MAIC, matching-adjusted indirect comparison; QALY, quality-adjusted life-year

B.3.8. Sensitivity analyses

B.3.8.1 Deterministic sensitivity analysis

The deterministic sensitivity analysis (DSA) is based on the modification of basic clinical and economic assumptions in the model, to test the strength of the conclusions of the analysis over a range of assumed input values.

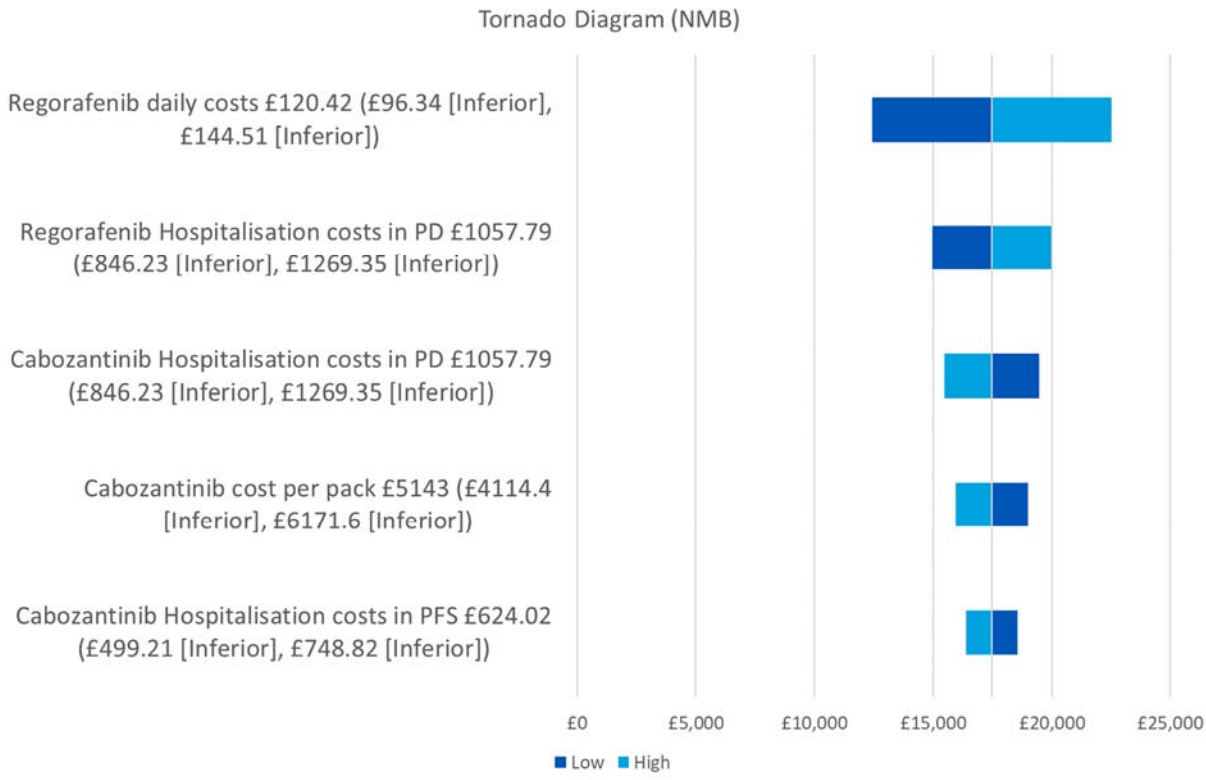
The analysis was performed in a structured manner on an exhaustive list of parameters (including costs, response to treatment, safety and efficacy, and utilities),

which involves varying an individual parameter through a range of plausible values (e.g. a low and a high estimate) whilst holding other parameters fixed and assessing the effect on the overall outcome.

The DSA was performed for input parameters of the model within their 95% confidence interval or their most plausible ranges. If no information was available a range of ± 25% from the point estimate is assumed. The NMB results of the DSA, using a WTP of £30,000 per QALY, for the most influential parameters are displayed for each of the key modelling scenarios. Results include the cabozantinib PAS price.

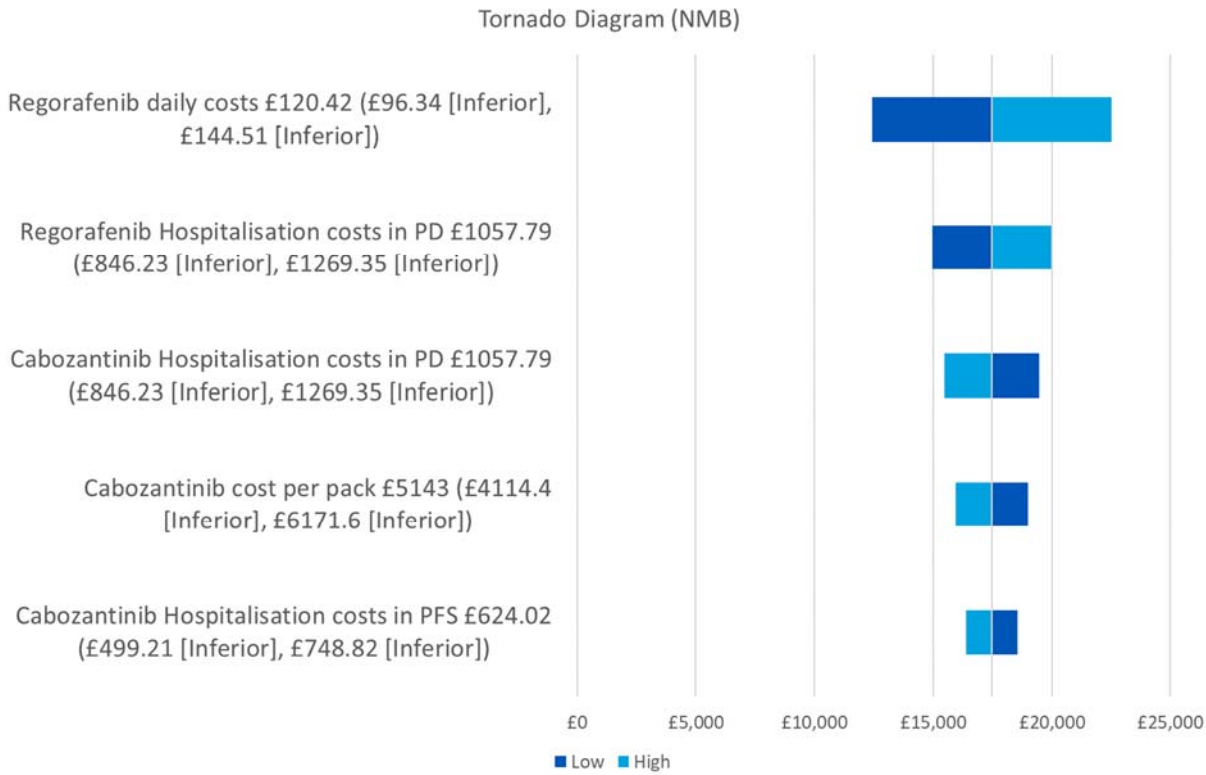
The results of the DSA for the anchored MAIC, constant HR scenario are shown in Figure 42. The results of the DSA for the anchored MAIC, time-varying HR scenario are shown in Figure 43. The results of the DSA are shown in Figure 44 for the unanchored MAIC scenario. The key drivers were the drug acquisition costs for cabozantinib and regorafenib as well as varying the health state cost between treatments.

Figure 42: Anchored MAIC, constant HR scenario tornado diagram



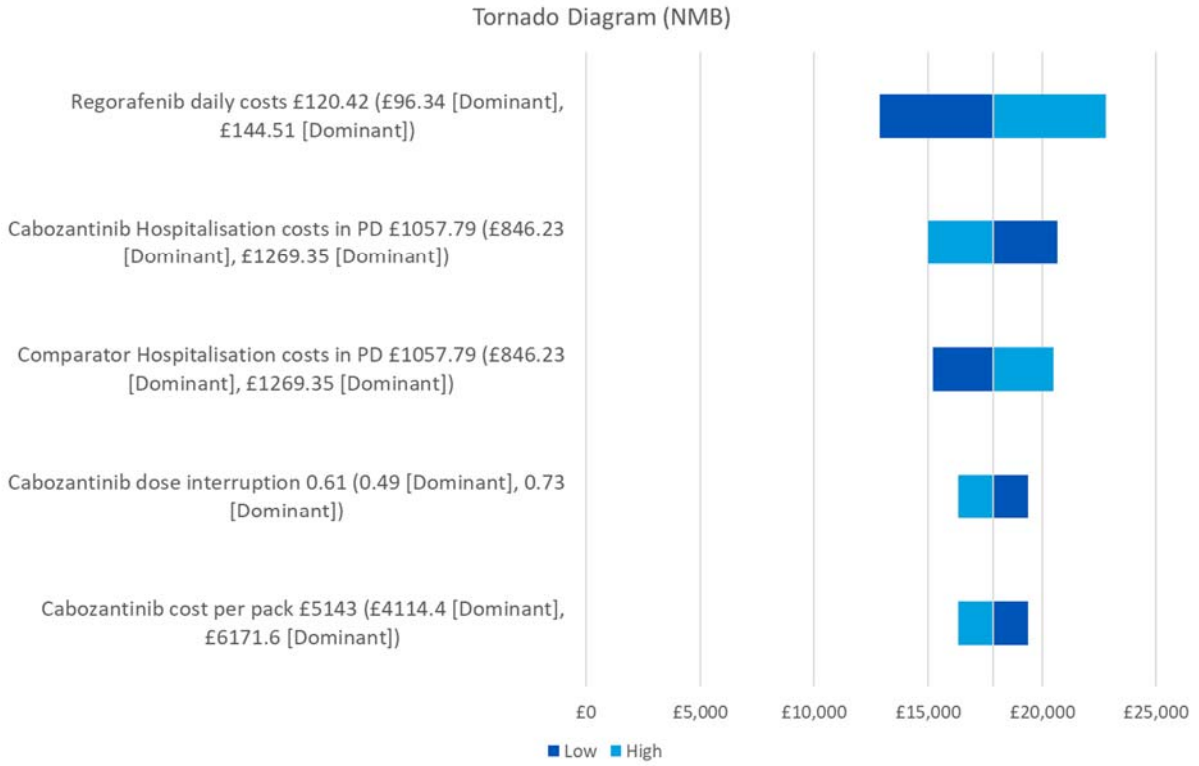
Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; NMB, net monetary benefit; PD, progressed disease; PFS, progression-free survival

Figure 43: Anchored MAIC, time-varying HR scenario tornado diagram



Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; NMB, net monetary benefit; PD, progressed disease; PFS, progression-free survival

Figure 44: Unanchored MAIC scenario tornado diagram



Abbreviations: MAIC, matching-adjusted indirect comparison; NMB, net monetary benefit; PD, progressed disease; PFS, progression-free survival

B.3.8.2 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed to account for multivariate and stochastic uncertainty in the model. 1,000 simulations were run.

The PSA was conducted to simultaneously take into account the uncertainty associated with parameter values. The implementation of PSA involved assigning specific parametric distributions and repeatedly sampling mean parameter values. Sampling was based on parameter distribution around the mean estimate at a 95% confidence interval, constructed using reported standard errors where available. A default margin of error of 20% around the mean estimate was applied where standard errors of the mean were not available/ not reported. The distributions used for the type of variable is shown in Table 58.

Table 58. Probabilistic sensitivity analysis distributions

Variable	Distribution
Cost; disutilities	Gamma

Variable	Distribution
Probabilities, utilities	Beta
Survival parameters	Multivariate normal
Hazard ratios, odd ratios	Lognormal

The probabilistic results for each of the key modelling scenarios are shown in Table 59. Results include the cabozantinib PAS price. The probabilistic ICERs were also unstable given the small incremental QALYs.

Table 59. Probabilistic results

Technologies	Total costs	Total LYs	Total QALYs	ICER (£/QALY)
Anchored MAIC, constant HR				
Cabozantinib	[redacted]	1.40	[redacted]	
Regorafenib	£54,628	1.49	1.05	
Incremental	[redacted]	-0.10	[redacted]	SW £269,333
Anchored MAIC, time-varying HR				
Cabozantinib	[redacted]	1.70	[redacted]	
Regorafenib	£60,484	1.81	1.27	
Incremental	[redacted]	1.69	[redacted]	SW £229,658
Unanchored MAIC				
Cabozantinib	[redacted]	1.69	[redacted]	
Regorafenib	£55,424	1.53	1.07	
Incremental	[redacted]	0.16	[redacted]	Dominant

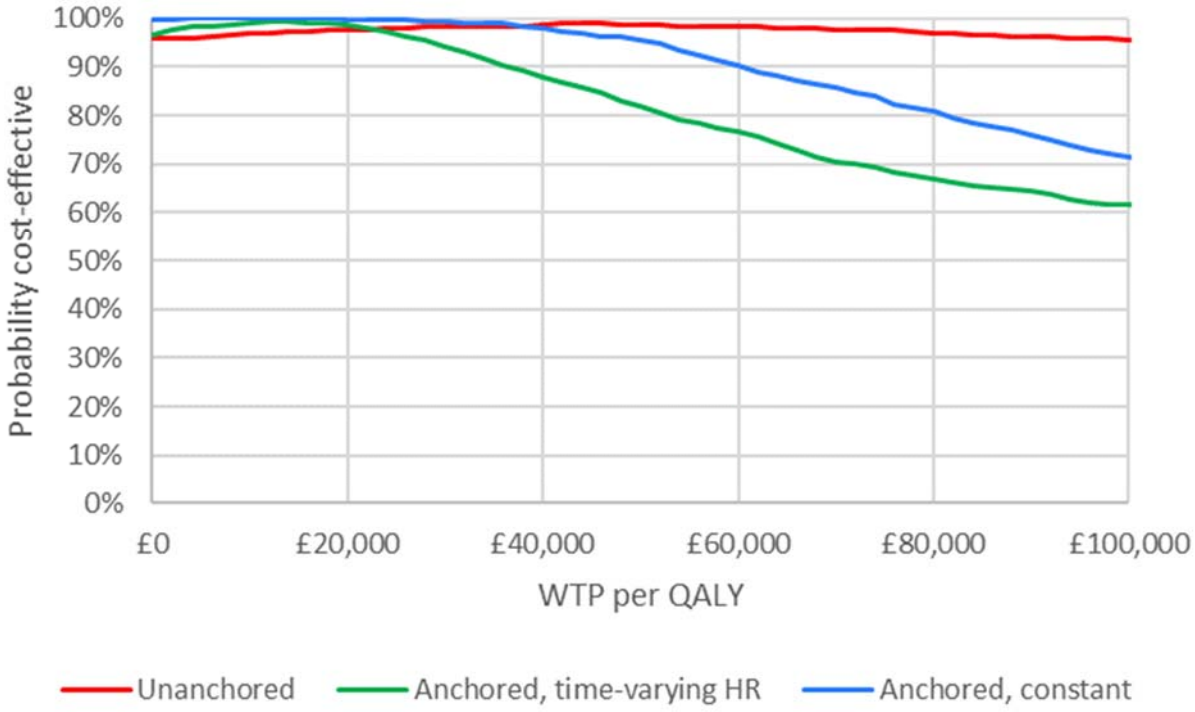
Abbreviations: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; LY, life-year; MAIC, matching-adjusted indirect comparison; QALY, quality-adjusted life-year; SW, south west

The cost-effectiveness plane and cost-effectiveness acceptability curve (CEAC) are shown in Figure 45 and Figure 46, respectively. The simulations were centered around 0 incremental QALYs and were mostly cost saving; therefore, the probability of being cost-effective at £0 WTP per QALY was above 90%. At a WTP of £20,000 and £30,000 per QALY, the probability of being cost-effective remains above 90% for all scenarios. The unanchored MAIC scenario had a probability of 78% for a positive incremental QALY. The probability in the anchored MAIC scenarios was 29% and 43% for the constant HR and time-varying HR respectively.

Figure 45: Probabilistic sensitivity analysis cost-effectiveness plane

Abbreviations: HR, hazard ratio; QALY, quality-adjusted life-year

Figure 46: Probabilistic sensitivity analysis cost-effectiveness acceptability curve



Abbreviations: HR, hazard ratio; QALY, quality-adjusted life-year; WTP, willingness to pay

B.3.8.3 Scenario analysis

The list of scenarios explored in the model are listed in Table 60 and the results are included in Table 61. Across all scenarios, the ICERs were unstable from the small incremental QALYs. The Bucher ITC and TTD scenarios do not adjust for population differences and so must be interpreted with caution.

Table 60. Scenario analyses explored in the model

Model setting	Base case	Scenario analysis	Justification
Time horizon	15 years	10 years	A shorter time horizon was modelled to test the impact on costs and outcomes over time
Treatment duration	TTD equal to PFS	TTD curve Cabozantinib: lognormal Regorafenib: log-logistic	The TTD was used to test the impact of treating beyond progression
RDI	Cabozantinib: 61.0% Regorafenib: 90.1%	Cabozantinib: 100% Regorafenib: 100%	Full pack dosing was tested to see the most conservative scenario for costs
Discount rate for costs and outcomes	3.5%	0%	
		6%	
Drug acquisition costs	Cabozantinib: PAS price Regorafenib: list price	Cabozantinib: list price Regorafenib: list price	
OS & PFS extrapolations	Anchored MAIC, time-varying HR: OS: log-logistic PFS: log-logistic Unanchored MAIC: Cabozantinib OS: log-logistic Cabozantinib PFS: Generalised gamma Regorafenib OS: log-logistic Regorafenib PFS: Generalised gamma	Anchored MAIC, time-varying HR: OS: lognormal PFS: lognormal Unanchored MAIC: Cabozantinib OS: lognormal Cabozantinib PFS: lognormal Regorafenib OS: lognormal Regorafenib PFS: lognormal	The second best fitting was tested to see the impact of model choice
ITC	MAIC	Bucher ITC using CELESTIAL ITT cabozantinib OS (Weibull), PFS (Weibull) and utility	The Bucher methodology was tested to assess the results when trial randomisation was not broken, however no

Model setting	Base case	Scenario analysis	Justification
			population adjustment is taken into account
Wastage	No additional wastage cost	Quarter pack of wastage cost	An alternative wastage assumption from TA474 was tested
Utility	CELESTIAL 2L PF: [redacted] PD: [redacted] AE: [redacted]	RESORCE PF: 0.811 PD: -0.048 AE: -0.014	An alternative utility source was tested

Abbreviations: AE, adverse event; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; OS, overall survival; PAS, patient access scheme; PD, progressed disease; PF, progression-free; PFS, progression-free survival; RDI, relative dose intensity; TTD, time to treatment discontinuation

Table 61. Scenario analyses results

Scenario	Incr. costs	Incr. QALYs	ICER (£/QALY)
Base case			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £290,383
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £300,170
Unanchored MAIC	[redacted]	[redacted]	Dominant
Time horizon			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £290,487
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £326,671
Unanchored MAIC	[redacted]	[redacted]	Dominant
Treatment duration			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £385,422
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £490,219
Unanchored MAIC	[redacted]	[redacted]	Dominant
RDI			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £259,254
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £253,353
Unanchored MAIC	[redacted]	[redacted]	Dominant
Discount rate for costs and outcomes – 0%			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £275,360
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £269,052
Unanchored MAIC	[redacted]	[redacted]	Dominant
Discount rate for costs and outcomes – 6%			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £300,927
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £321,673
Unanchored MAIC	[redacted]	[redacted]	Dominant
Drug acquisition costs			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £25,227
Anchored MAIC, time-varying HR	[redacted]	[redacted]	Dominated
Unanchored MAIC	[redacted]	[redacted]	NE £30,255
OS & PFS extrapolations			

Scenario	Incr. costs	Incr. QALYs	ICER (£/QALY)
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £252,289
Unanchored MAIC	[redacted]	[redacted]	Dominant
ITC			
Bucher ITC	[redacted]	[redacted]	SW £162,411
Wastage			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £298,582
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £309,195
Unanchored MAIC	[redacted]	[redacted]	Dominant
Utility			
Anchored MAIC, constant HR	[redacted]	[redacted]	SW £290,745
Anchored MAIC, time-varying HR	[redacted]	[redacted]	SW £321,707
Unanchored MAIC	[redacted]	[redacted]	Dominant

Abbreviations: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; NE, north east; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life-year; RDI, relative dose intensity; SW, south west

B.3.8.4 Summary of sensitivity analyses results

- The DSA showed that the key drivers were the drug acquisition costs of cabozantinib and regorafenib as well as varying the health state costs between treatments. The impact of AE cost and disutility were low.

- The PSA showed that at the discounted price for cabozantinib, most simulations were cost saving. The incremental benefit was centred around 0, supporting an equal efficacy assumption between cabozantinib and regorafenib.

- The scenario analysis showed that testing the assumptions in the model did not have a substantial impact on the incremental benefit. When allowing treatment beyond progression, the incremental costs decreased.

B.3.9. Subgroup analysis

No subgroup analyses were considered as part of this submission.

B.3.10. Validation

A clinical expert advisory board was consulted to validate the model inputs and assumptions (3). Once the model was finalised, it was validated by internal modellers. A programmer (other than the one that built the model) reviewed all formulae and labelling in the model.

B.3.11. Interpretation and conclusions of economic evidence

The CCA demonstrated in the FTA that, when equivalent clinical effectiveness is assumed, cabozantinib is cost-saving when compared to regorafenib using the cabozantinib PAS price. To assess the uncertainty in the equal efficacy assumption a cost-effectiveness analysis was conducted by modelling the scenarios from the ITC. All of the ITC scenarios were cost-effective when using the cabozantinib PAS.

There was evidence to suggest that the ITC could not account for the heterogeneity between the CELESTIAL and RESORCE trials such as patient tolerability to previous sorafenib treatment. Therefore, the 3 key approaches to modelling the ITC may underestimate the efficacy of cabozantinib. The anchored MAIC approach may theoretically benefit by accounting for prognostic factors through the relative effect to the placebo arm. Therefore, a relative measure of effect was explored to model the anchored approach. The constant HR scenario produced a point estimate incremental QALY of [redacted]; however, the proportional hazards assumption was violated between the trials and so the use of a constant HR was not appropriate. A time-varying HR was explored; however, this introduced additional uncertainty from modelling and extrapolating the hazards of the two placebo arms. This scenario generated a similar point estimate to the constant HR scenario for the incremental QALY [redacted] vs [redacted]). Given these limitations, the unanchored MAIC approach was also explored and this scenario resulted in a point estimate incremental QALY of [redacted]. The unanchored MAIC aligns with the independent findings from RWE such as the RWE MAIC study by Casadei-Gardini et al (2021) which assessed a population for regorafenib that more closely resembled the cabozantinib population with respect to sorafenib tolerability, a variable that was not possible to adjust for in the ITC used in the submission. In conclusion, the small QALY difference between treatments and the additional benefit in the unanchored MAIC which is validated from RWE, would indicate that cabozantinib has similar or greater benefit than regorafenib.

The uncertainty in each of the key ITC approaches were assessed in the PSA. The simulations showed that the incremental benefit did not favour either treatment as the simulations were centred around 0 incremental QALYs. The scenario analysis showed that alternative assumptions had little impact on the range of incremental QALYs.

The DSA showed that the model is sensitive to the drug acquisition cost of cabozantinib and regorafenib. The 100% RDI scenario showed that in the most conservative assumption for wastage in clinical practice, the cost saving of cabozantinib is reduced. The TTD scenario showed that there was greater cost savings if patients are treated beyond progression as the regorafenib duration of treatment in RESORCE was greater than cabozantinib. The efficacy for treatment beyond progression could not be adjusted for in the treat to progression scenarios. It is possible this may have introduced a bias towards an improvement for regorafenib in terms of its OS endpoint as evaluated in the RESORCE trial.

Given the sum of the evidence from the different ITC scenarios, assessment of uncertainty, clinical expert validation and RWE, cabozantinib offers the NHS an equally efficacious, cost-saving and tolerable alternative to regorafenib treatment of adults with advanced HCC who have previously been treated with sorafenib.

B.4. References

1. National Institute for Health and Care Excellence. Regorafenib for previously treated advanced hepatocellular carcinoma TA555. 2019.
2. National Institute for Health and Care Excellence. Atezolizumab with bevacizumab for treating advanced or unresectable hepatocellular carcinoma TA666. 2020.
3. Ipsen. Advisory board report, data on file. 2021.
4. National Health Service England. National Cancer Drugs Fund List. ver1.200 December 2021 [Available from: https://www.england.nhs.uk/wp-content/uploads/2017/04/NationalCDF_List_ver1.200_-22122021_MASTER.pdf.
5. National Institute for Health and Care Excellence. Single Technology Appraisal (STA). Cabozantinib for treating advanced hepatocellular carcinoma after prior therapy ID1243. Response to consultee and commentator comments on the draft remit and draft scope (pre-referral) 2018 [Available from: https://www.nice.org.uk/guidance/ta582/documents/scope-consultation-comments-and-responses.
6. 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;379(1):54-63.
7. Xiang Q CW, Ren M, Wang J, Zhang H, Deng DY, et al. Cabozantinib suppresses tumor growth and metastasis in hepatocellular carcinoma by a dual blockade of VEGFR2 and MET. Clinical cancer research : an official journal of the American Association for Cancer Research. 2014;20(11):2959-70.
8. Yakes FM CJ, 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. Molecular cancer therapeutics. 2011;10(12):2298-308.
9. Bentzien F, Zuzow M, Heald N, Gibson A, Shi Y, Goon L, et al. In vitro and in vivo activity of cabozantinib (XL184), an inhibitor of RET, MET, and VEGFR2, in a model of medullary thyroid cancer. Thyroid : official journal of the American Thyroid Association. 2013;23(12):1569-77.
10. Escudier B, Lougheed JC, Albiges L. Cabozantinib for the treatment of renal cell carcinoma. Expert Opin Pharmacother. 2016;17(18):2499-504.
11. Kurzrock R, Sherman SI, Ball DW, Forastiere AA, Cohen RB, Mehra R, et al. Activity of XL184 (Cabozantinib), an oral tyrosine kinase inhibitor, in patients with medullary thyroid cancer. J Clin Oncol. 2011;29(19):2660-6.
12. Kelley RK, Verslype C, Cohn AL, Yang TS, Su WC, Burris H, et al. Cabozantinib in hepatocellular carcinoma: results of a phase 2 placebo-controlled randomized discontinuation study. Ann Oncol. 2017;28(3):528-34.
13. Joint Formulary Committee. British National Formulary (online) London: BMJ Group and Pharmaceutical Press <http://www.medicinescomplete.com> [Accessed on [01/12/2021]] 2022 [
14. Orphanet. Hepatocellular carcinoma. [15 June 2018]; Available from: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=EN&Expert=88673.
15. Wu J. The changing epidemiology of hepatocellular carcinoma in Asia versus United States and Europe. . Adv Mod Oncol Res. 2017;3:p. 51-8.

16. Blachier M, Leleu H, Peck-Radosavljevic M, Valla DC, Roudot-Thoraval F. The burden of liver disease in Europe: a review of available epidemiological data. J Hepatol. 2013;58(3):593-608.
17. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology. 2012;142(6):1264-73.e1.
18. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011;61(2):69-90.
19. McGlynn KA, Petrick JL, London WT. Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability. Clin Liver Dis. 2015;19(2):223-38.
20. Velázquez RF, Rodríguez M, Navascués CA, Linares A, Pérez R, Sotorríos NG, et al. Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis. Hepatology. 2003;37(3):520-7.
21. Burton A, Tataru D, Driver RJ, Bird TG, Huws D, Wallace D, et al. Primary liver cancer in the UK: Incidence, incidence-based mortality, and survival by subtype, sex, and nation. JHEP Rep. 2021;3(2):100232.
22. Cancer Research UK. Liver Cancer incidence statistics. [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/incidence#heading-Zero.
23. Cancer Research UK. Liver cancer mortality statistics. [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/mortality#heading-Zero.
24. EASL. EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol. 2012;56(4):908-43.
25. Vogel A, Cervantes A, Chau I, Daniele B, Llovet JM, Meyer T, et al. Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018;29(Suppl 4):iv238-iv55.
26. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology. 2011;53(3):1020-2.
27. Durand F, Valla D. Assessment of prognosis of cirrhosis. Semin Liver Dis. 2008;28(1):110-22.
28. EASL. Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol. 2018;69(1):182-236.
29. Shiina S, Tateishi R, Imamura M, Teratani T, Koike Y, Sato S, et al. Percutaneous ethanol injection for hepatocellular carcinoma: 20-year outcome and prognostic factors. Liver Int. 2012;32(9):1434-42.
30. Shah SA, Cleary SP, Wei AC, Yang I, Taylor BR, Hemming AW, et al. Recurrence after liver resection for hepatocellular carcinoma: risk factors, treatment, and outcomes. Surgery. 2007;141(3):330-9.
31. Bruix J, Reig M, Sherman M. Evidence-Based Diagnosis, Staging, and Treatment of Patients With Hepatocellular Carcinoma. Gastroenterology. 2016;150(4):835-53.
32. Park JW, Chen M, Colombo M, Roberts LR, Schwartz M, Chen PJ, et al. Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. Liver Int. 2015;35(9):2155-66.
33. Verslype C, Rosmorduc O, Rougier P. Hepatocellular carcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2012;23 Suppl 7:vii41-8.

34. Ryder SD. British Society of Gastroenterology: Guidelines for the diagnosis and treatment of hepatocellular carcinoma (HCC) in adults. Gut. 2003;52:iii1–iii8.
35. ESMO. Hepatocellular carcinoma: European Society for Medical Oncology (ESMO) Clinical Practice Guidelines for diagnosis, treatment and follow-up (March 2021 update): . 2021.
36. Bruix J, Chan SL, Galle PR, Rimassa L, Sangro B. Systemic treatment of hepatocellular carcinoma: An EASL position paper. J Hepatol. 2021;75(4):960-74.
37. ILCA Systemic Therapy Guidance 2020 [Available from: https://livercentral.org/wp-content/uploads/2021/12/Systemic-therapy-guidelines-V1.2.pdf.
38. NICE. Sorafenib for treating advanced hepatocellular carcinoma TA474. 2017.
39. NICE. Lenvatinib for untreated advanced hepatocellular carcinoma TA551. 2018.
40. Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2017;389(10064):56-66.
41. Pinato DJ, Brown MW, Trousil S, Aboagye EO, Beaumont J, Zhang H, et al. Integrated analysis of multiple receptor tyrosine kinases identifies Axl as a therapeutic target and mediator of resistance to sorafenib in hepatocellular carcinoma. British Journal of Cancer. 2019;120(5):512-21.
42. Xiang Q-F, Zhan M-X, Li Y, Liang H, Hu C, Huang Y-M, et al. Activation of MET promotes resistance to sorafenib in hepatocellular carcinoma cells via the AKT/ERK1/2-EGR1 pathway. Artificial Cells, Nanomedicine, and Biotechnology. 2019;47(1):83-9.
43. Firtina Karagonlar Z, Koc D, Iscan E, Erdal E, Atabey N. Elevated hepatocyte growth factor expression as an autocrine c-Met activation mechanism in acquired resistance to sorafenib in hepatocellular carcinoma cells. Cancer Science. 2016;107(4):407-16.
44. Exelixis. CLINICAL STUDY REPORT XL184-309: A phase 3, randomized, double-blind, controlled study of cabozantinib (XL184) vs placebo in subjects with hepatocellular carcinoma who have received prior sorafenib. 2018.
45. King J, Palmer DH, Johnson P, Ross P, Hubner RA, Sumpter K, et al. Sorafenib for the Treatment of Advanced Hepatocellular Cancer - a UK Audit. Clin Oncol (R Coll Radiol). 2017;29(4):256-62.
46. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007;5:70.
47. Llovet JM, Decaens T, Raoul JL, Boucher E, Kudo M, Chang C, et al. Brivanib in patients with advanced hepatocellular carcinoma who were intolerant to sorafenib or for whom sorafenib failed: results from the randomized phase III BRISK-PS study. J Clin Oncol. 2013;31(28):3509-16.
48. Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. J Clin Epidemiol. 2005;58(12):1217-9.
49. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof. 2005;28(2):172-91.

50. Bucher HC GG, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997;50(6):683-91.:683-91.
51. Finn RS, Merle P, Granito A, Huang YH, Bodoky G, Pracht M, et al. Outcomes of sequential treatment with sorafenib followed by regorafenib for HCC: Additional analyses from the phase III RESORCE trial. J Hepatol. 2018;69(2):353-8.
52. Waldschmidt D, Granito A, Merle P, Huang Y, Bodoky G, Yokosuka O, et al. Overall Survival (OS) Update: 2-year follow-up from the phase-3 RESORCE trial of Regorafenib for patients with hepatocellular carcinoma (HCC) progressing on Sorafenib. Zeitschrift für Gastroenterologie. 2019;57(09):KV 214.
53. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Medical Research Methodology. 2012;12(1):9.
54. Signorovitch JE, Wu EQ, Yu AP, Gerrits CM, Kantor E, Bao Y, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. Pharmacoeconomics. 2010;28(10):935-45.
55. Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012;15(6):940-7.
56. Nash P, McInnes IB, Mease PJ, Thom H, Hunger M, Karabis A, et al. Secukinumab Versus Adalimumab for Psoriatic Arthritis: Comparative Effectiveness up to 48 Weeks Using a Matching-Adjusted Indirect Comparison. Rheumatology and therapy. 2018;5(1):99-122.
57. Ipsen. Advisory board report, data on file. 2018.
58. Venables WNaR, B.D. Modern Applied Statistics with S. Fourth Edition2002.
59. White H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica. 1980;48(4):817-38.
60. Woodruff RS. Confidence Intervals for Medians and Other Position Measures. Journal of the American Statistical Association. 1952;47(260):635-46.
61. Mandel M. Simulation-Based Confidence Intervals for Functions With Complicated Derivatives. The American Statistician. 2013;67(2):76-81.
62. Kelley RK, Mollon P, Blanc JF, Daniele B, Yau T, Cheng AL, et al. Comparative Efficacy of Cabozantinib and Regorafenib for Advanced Hepatocellular Carcinoma. Adv Ther. 2020;37(6):2678-95.
63. Casadei-Gardini A, Rimassa L, Rimini M, Yoo C, Ryoo BY, Lonardi S, et al. Regorafenib versus cabozantinb as second-line treatment after sorafenib for unresectable hepatocellular carcinoma: matching-adjusted indirect comparison analysis. J Cancer Res Clin Oncol. 2021;147(12):3665-71.
64. Wang D, Yang X, Lin J, Bai Y, Long J, Yang X, et al. Comparing the efficacy and safety of second-line therapies for advanced hepatocellular carcinoma: a network meta-analysis of phase III trials. Therap Adv Gastroenterol. 2020;13:1756284820932483.
65. Bakouny Z, Assi T, El Rassy E, Nasr F. Second-line Treatments of Advanced Hepatocellular Carcinoma: Systematic Review and Network Meta-Analysis of Randomized Controlled Trials. J Clin Gastroenterol. 2019;53(4):251-61.
66. Sonbol MB, Riaz IB, Naqvi SAA, Almquist DR, Mina S, Almasri J, et al. Systemic Therapy and Sequencing Options in Advanced Hepatocellular Carcinoma:

A Systematic Review and Network Meta-analysis. JAMA Oncology. 2020;6(12):e204930-e.
67. Park R, Lopes da Silva L, Nissaisorakarn V, Riano I, Williamson S, Sun W, et al. Comparison of Efficacy of Systemic Therapies in Advanced Hepatocellular Carcinoma: Updated Systematic Review and Frequentist Network Meta-Analysis of Randomized Controlled Trials. J Hepatocell Carcinoma. 2021;8:145-54.
68. Tovoli F, Dadduzio V, De Lorenzo S, Rimassa L, Masi G, Iavarone M, et al. Real-Life Clinical Data of Cabozantinib for Unresectable Hepatocellular Carcinoma. Liver Cancer. 2021;10(4):370-9.
69. Finkelmeier F, Scheiner B, Leyh C, Best J, Fründt TW, Czauderna C, et al. Cabozantinib in Advanced Hepatocellular Carcinoma: Efficacy and Safety Data from an International Multicenter Real-Life Cohort. Liver Cancer. 2021;10(4):360-9.
70. Parikh ND, Singal AG, Hutton DW. Cost effectiveness of regorafenib as second-line therapy for patients with advanced hepatocellular carcinoma. Cancer. 2017;123(19):3725-31.
71. Soto-Perez-de-Celis E, Aguiar PN, Cordón ML, Chavarri-Guerra Y, de Lima Lopes GJJotNCCN. Cost-effectiveness of cabozantinib in the second-line treatment of advanced hepatocellular carcinoma. 2019;17(6):669-75.
72. Liao W, Huang J, Hutton D, Zhu G, Wu Q, Wen F, et al. Cost-effectiveness analysis of cabozantinib as second-line therapy in advanced hepatocellular carcinoma. 2019;39(12):2408-16.
73. Shlomai A, Leshno M, Goldstein DAJTaig. Cabozantinib for patients with advanced hepatocellular carcinoma: a cost-effectiveness analysis. 2019;12:1756284819878304.
74. Shlomai A, Leshno M, Goldstein DAJPO. Regorafenib treatment for patients with hepatocellular carcinoma who progressed on sorafenib—A cost-effectiveness analysis. 2018;13(11):e0207132.
75. Upadhyay N, Mistry R, Suri G, Phelps H, Guest S, Muston DJViH. PCN96 US COST-EFFECTIVENESS OF PEMBROLIZUMAB COMPARED TO REGORAFENIB AND CABOZANTINIB IN SECOND LINE TREATMENT OF ADVANCED HEPATOCELLULAR CARCINOMA. 2019;22:S73-S4.
76. Kim J, Hay JJViH. Cost-Effectiveness Analysis of Regorafenib in Patients with Advanced Hepatocellular Carcinoma. 2018;21:S29.
77. Sieg M, Hartmann M, Settmacher U, Arefian HJBg. Comparative cost-effectiveness of cabozantinib as second-line therapy for patients with advanced hepatocellular carcinoma in Germany and the United States. 2020;20:1-12.
78. National Institute for Health and Care Excellence. Guide to the processes of technology appraisal 2018 [Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/technology-appraisal-processes-guide-apr-2018.pdf.
79. National Institute for Health and Care Excellence. What we do 2018 [Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf
80. Latimer N. NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA 2011.
81. Ipsen. Parametric survival analyses of CELESTIAL ITT, data on file. 2021.

82. Li YMS, Mollon P, Guerra I. Healthcare resource use among patients with advanced hepatocellular carcinoma (aHCC) in the UK. HCC-UK Annual Conference. 2020.
83. National Health Service. Reference costs 2019/20. 2021.
84. Personal Social Service Unit. Unit costs of health and social care. 2021.
85. Ipsen. Clinical expert interviews. 2022.
86. Coyle D SN, Ashworth A, et al. Costs of palliative care in the community, in hospitals and in hospices in the UK. Crit Rev Oncol Hematol. 1999;32(2):71-85.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Clarification questions

February 2022

File name	Version	Contains confidential information	Date
ID3917 cabozantinib clarification letter to PM for company [CIC]	1.0	Yes	16.03.2022

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Decision problem and target population

A1. Priority. Company submission (CS), Section B.1.1, page 9 and Table 1, page 12. The CS states that *"The submission covers the technology's full marketing authorisation for this indication."* In addition, CS Table 1 indicates that the decision problem addressed in the CS is the same as the final NICE scope. Please clarify the intended target population for cabozantinib in light of the following issues:

(i) The NICE recommendation for regorafenib is restricted to patients with Child Pugh A and Eastern Cooperative Oncology Group (ECOG) Performance Status (PS) of 0 or 1 and the CS (page 24) states that *"It is proposed that cabozantinib is positioned where regorafenib is currently used in practice."* This is narrower than the marketing authorisation for cabozantinib.

(ii) The CS (page 80) acknowledges that the indirect treatment comparison using the Bucher approach is subject to bias, whilst the matching-adjusted indirect comparison (MAIC) is restricted to the second-line populations of CELESTIAL and RESORCE. The population reflected in the MAIC is narrower than the marketing authorisation for cabozantinib.

Company response

The CELESTIAL trial only included one patient (<1%) in the cabozantinib arm with Child Pugh B whilst the RESORCE trial had five patients (1%) in the regorafenib arm with Child Pugh B. As we propose to displace regorafenib use in HCC, we assumed that the NICE recommendation for cabozantinib in terms of Child Pugh status would be the same as regorafenib. However, we would support recommendation in a broader population as per its licensed indication without restriction based on Child Pugh status. Similarly, the regorafenib licensed indication also makes no mention of Child Pugh status in the wording of the indication; only NICE places a reimbursement restriction based on Child Pugh status for regorafenib.

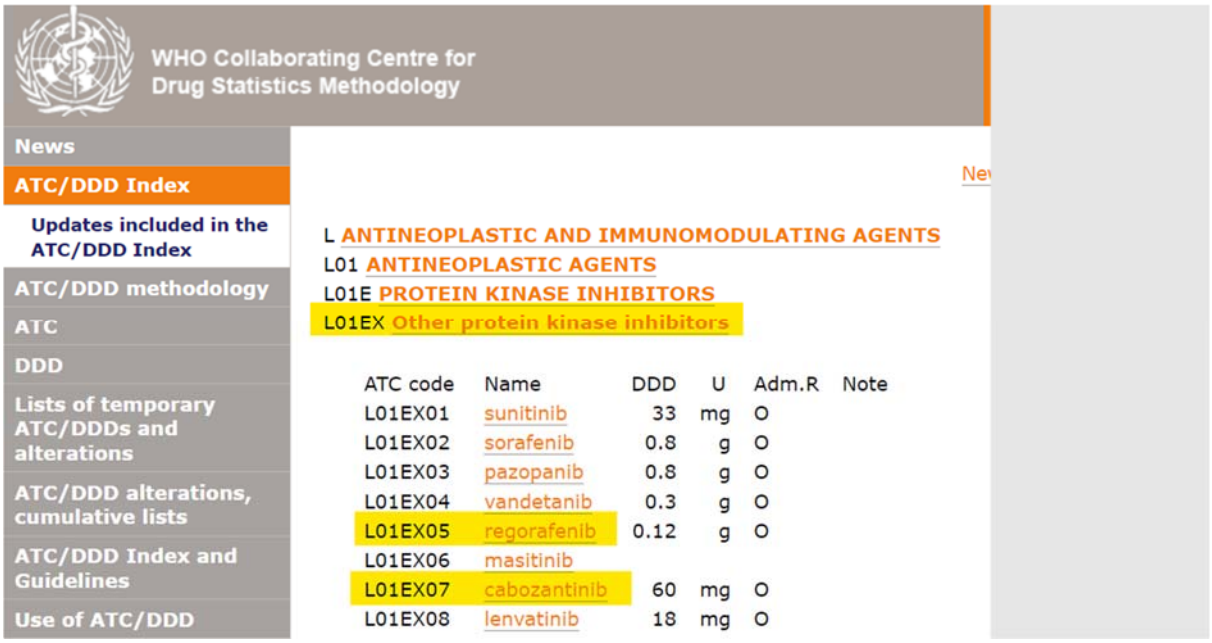
Within the CELESTIAL trial there was a mixed second and third-line population, whereas the RESORCE trial only included a second-line population with patients who tolerated sorafenib. The availability of IPD for CELESTIAL enabled the isolation of a pure second-line subpopulation to be compared with RESORCE to produce the MAIC. In the CELESTIAL trial, there were 130 (28%) patients in the cabozantinib arm that received two prior regimens for advanced HCC, with a further 2 patients receiving 3 prior treatments. The HRs for these two prior regimen subgroups were 0.90 (0.63-1.29 [95% CI]) and 0.58 (0.41-0.83 [95% CI]), for OS and PFS, respectively. We therefore would support a recommendation within this third-line subgroup; however, a MAIC could only be conducted for the second-line population because of the RESORCE population restrictions and hence the population reflected in the MAIC is narrower than the marketing authorisation for cabozantinib.

A2. CS, Section B.1.1, page 11. The CS states that *"Cabozantinib and regorafenib belong to the same drug class of tyrosine kinase inhibitors (TKIs)."* Please outline the specific similarities and differences between cabozantinib and regorafenib in terms of mechanism of action and targeting of specific receptor tyrosine kinases.

Company response

Both cabozantinib and regorafenib are tyrosine kinase inhibitors that are grouped under the L01EX category (other protein kinase inhibitors) according to WHO ATC, as shown below in Figure 1.

Figure 1: WHO Anatomical Therapeutic Chemical (ATC) coding for Protein Kinase Inhibitors



Although they share a drug class and are orally administered, cabozantinib and regorafenib do have some differences in their molecular targeting profiles.

Regorafenib targets multiple receptor tyrosine kinases (RTKs), including those involved in tumour angiogenesis (VEGFR-1, -2, -3, TIE2), oncogenesis (KIT, RET, RAF-1, BRAF, BRAFV600E), metastasis (VEGFR3, PDGFR, FGFR) and tumour immunity (CSF1R) [Bayer Plc, 2022]. Cabozantinib also inhibits multiple RTKs, including those implicated in tumour angiogenesis (VEGFR-1, -2, -3, TIE-2) and oncogenesis (RET), but it additionally targets MET receptor, involved in tumour growth and invasion and the MET receptor involved in modulation of tumour immunity. High expression of MET or AXL may be associated with poor prognosis in patients with hepatocellular carcinoma, (Ueki et al. 1997, Liu et al. 2016, Santoni et al. 2021) and increased MET expression or activation has been associated with previous sorafenib treatment in patients with hepatocellular carcinoma and with sorafenib resistance in preclinical models (Rimassa et al. 2016, Rimassa et al. 2018, Xiang et al. 2014, Firtina et al. 2016) and thus cabozantinib has a biologically plausible rationale to treat patients who are resistant to sorafenib.

Evidence searches

A3. CS, Appendix D, Section D.1.1.4, search strategy. Searches were carried out on the 27th March 2018 followed by an update on the 22nd February 2021 and January 2022. Please explain why the updated clinical search strategy in Embase in Table 4 (page 15) and Table 9 (page 21) differ (shorter search strings) from the original search strategy in Table 1 (page 12). Please explain what impact this will have had on search recall and on the subsequent findings of the review.

Company response

We discussed with the ERG and NICE at the Decision Problem meeting that following a delay to the submission by seven months due to NICE capacity issues how we would create a more focused clinical search strategy, with the advice from the ERG to include evidence in the SLR that no updates have been made to the literature since April 2021. We did this in a pragmatic way using shorter search strings and do not believe any relevant data has been missed in doing this.

Clinical evidence for cabozantinib - CELESTIAL trial

A4. **Priority**. CS, Section B.3.6.2, Table 18, pages 51 to 52. Please explain the difference between the three measures of progression-free survival (PFS) for CELESTIAL (primary, PFS2, PFS3) and state which of these is closest to the definition of PFS used in RESORCE.

Company response

In addition to the primary analysis of PFS (PFS1) in the CELESTIAL trial, sensitivity analyses were undertaken (PFS2 and PFS3) that included defining additional clinical outcomes as events and evaluated the impact of informative censoring. The definitions were as follows (see also Table 14, Document B):

- PFS1 analysis: earlier of radiographic progression per RECIST 1.1 or death due to any reason.
- PFS2 analysis: the following events were considered to be PFS events – radiographic progression per RECIST 1.1, death due to any reason, systemic

or local liver-directed NPACT/radiation (other than to bone), tumour resection, treatment discontinuation due to clinical deterioration.

- PFS3 analysis: the following events were considered to be PFS events – radiographic progression per RECIST 1.1, death due to any reason, treatment discontinuation due to clinical deterioration.

In RESORCE, PFS was defined as:

- the 'time (days) from date of randomisation to date of disease progression (radiological or clinical) or death due to any cause, if death occurs before progression is documented.

Therefore, PFS3 could be considered similar to the RESORCE definition of PFS; however, there is an additional nuance, as clinical deterioration is not a synonym for disease progression and was not further defined. Therefore, the PFS1 definition in CELESTIAL could be considered similar to the RESORCE definition. Nevertheless, the HRs for PFS1 and PFS3 were both 0.44 (0.36-0.52, 95% CI and 0.37-0.54, 95% CI respectively – see Table 18, Document B) and therefore do not impact the comparison with regorafenib for PFS.

A5. CS, Section B.3.8.1, pages 55 to 56. The CS states *"Cabozantinib was generally well tolerated."* However, Table 20 suggests that 68% of patients experienced Grade 3 or 4 adverse events (AEs), 50% experienced serious AEs, 89% experienced AEs leading to dose modifications and 21% discontinued treatment due to an AE. Please clarify the statement in the CS.

Company response

The statement "cabozantinib was generally well tolerated" views cabozantinib in the context of other multi-targeted potent TKIs. There is a known safety profile associated with TKI use and this can be managed by dose reduction and interruption, in addition to symptom management. Additionally, our clinical expert interviews suggest patients are monitored early on after initiating treatment, so Grade 3 or 4 events rarely result in patients being hospitalised. We acknowledge that "well tolerated" may not seem appropriate relative to placebo.

A6. CS, Section B.3.8.1, Table 20, page 56. Please provide an amended version of Table 20 which includes treatment-related Grade 3 or 4 AEs by treatment group for CELESTIAL.

Company response

Table 1 - Amended Table 20, p56 in CS. The CELESTIAL trial: summary of safety data including treatment-related Grade 3/4 AEs (safety population)

Adverse Events	Cabozantinib n=467 n (%)	Placebo n=237 n (%)
Any AE (all grades)	460 (99)	219 (92)
Grade 3 or 4 AEs	316 (68)	86 (36)
Treatment-related AEs (Grade 3 or 4)	439 (94)	148 (62)
Diarrhoea	42 (9)	2 (0.8)
PPES	78 (17)	0
Fatigue	39 (8.4)	6 (2.5)
Decreased appetite	22 (4.7)	0
Hypertension	69 (15)	2 (0.8)
Nausea	7 (1.5)	0
Vomiting	1 (0.2)	2 (0.8)
Dysphonia	2 (0.4)	0
Asthenia	19 (4.1)	4 (1.7)
Aspartate aminotransferase increased	36 (7.7)	11 (4.6)
Mucosal inflammation	8 (1.7)	0
Stomatitis	8 (1.7)	0
Weight decreased	5 (1.1)	0
Alanine aminotransferase increased	16 (3.4)	3 (1.3)
Dysgeusia	0	0
Rash	2 (0.4)	0
SAEs	232 (50)	87 (37)
Treatment-related SAEs	82 (18)	14 (5.9)
Treatment-related Grade 5 AE	6 (1.3)	1 (0.4)
Deaths (at any time, excluding PD)	314 (67)	167 (70)
AE leading to dose modification	416 (89)	94 (40)
AE leading to discontinuation of study drug	96 (21)	10 (4.2)

Abbreviations: AE, adverse events.
Source: Abou-Alfa et al., 2018, Exelixis, 2018.

A7. CS, Tables 20 and 21, pages 56 to 57, and CS Appendix D, Table 16, pages 38 to 39. Please provide amended versions of CS Tables 20 and 21 which also include the equivalent safety data from RESORCE.

Company response

Table 2 - Amended Table 20, p56 in CS. CELESTIAL and RESORCE trials: Summary of Safety Data (safety population)

Adverse Events	Cabozantinib n=467 n (%)	Placebo n=237 n (%)	Regorafenib n=374 n (%)	Placebo n=194 n (%)
Any AE (all grades)	460 (99)	219 (92)	374 (100)	179 (93)
Grade 3 or 4 AEs	316 (68)	86 (36)	248 (66)	75 (38)
Treatment-related AEs	439 (94)	148 (62)	346 (93)	100 (52)
SAEs	232 (50)	87 (37)	166 (44)	90 (47)
Treatment-related SAEs	82 (18)	14 (5.9)	36 (10)	5 (3)
Treatment-related Grade 5 AE	6 (1.3)	1 (0.4)	7(2)	2 (1)
Deaths (at any time, excluding PD)	314 (67)	167 (70)	50 (13)	38 (20)
AE leading to dose modification	416 (89)	94 (40)	255 (68)	60 (31)
AE leading to discontinuation of study drug	96 (21)	10 (4.2)	93 (25)	37 (19)

Abbreviations: AE, adverse events.
Source: Abou-Alfa et al., 2018, Bruix 2017, Exelixis, 2018.

Table 3 - Amended Table 21, p57. AEs* (any grade) reported in ≥10% of patients in either treatment group for CELESTIAL and RESORCE.

Event	Cabozantinib (number of patients (%)			Placebo (number of patients (%)			Regorafenib (number of patients (%)			Placebo (number of patients (%)		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
Any AE	460 (99)	270 (58)	46 (10)	219 (92)	80 (34)	6 (3)	374 (100)	208 (56)	40 (11)	179 (93)	61 (32)	14 (7)
Diarrhoea	251 (54)	45 (10)	1 (<1)	44 (19)	4 (2)	0	155 (41)	12 (3)	0	29 (15)	0	0
Decreased appetite	225 (48)	27 (6)	0	43 (18)	1 (<1)	0	NR	NR	NR	NR	NR	NR
PPES	217 (46)	79 (17)	0	12 (5)	0	0	198 (53)	47 (13)	NA	15 (8)	1 (1)	NA
Fatigue	212 (45)	49 (10)	0	70 (30)	10 (4)	0	151 (40)	34 (9)	NA	61 (32)	9 (5)	NA
Nausea	147 (31)	10 (2)	0	42 (18)	4 (2)	0	64 (17)	2 (1)	NA	26 (13)	0	NA
Hypertension	137 (29)	73 (16)	1 (<1)	14 (6)	4 (2)	0	116 (31)	56 (15)	1 (<1)	12 (6)	9 (5)	0
Vomiting	121 (26)	2 (<1)	0	28 (12)	6 (3)	0	47 (13)	3 (1)	0	13 (7)	1 (1)	0
Increase in AST level	105 (22)	51 (11)	4 (1)	27 (11)	15 (6)	1 (<1)	92 (25)	37 (10)	4 (1)	38 (20)	19 (10)	3 (2)
Asthenia	102 (22)	31 (7)	1 (<1)	18 (8)	4 (2)	0	NR	NR	NR	NR	NR	NR
Dysphonia	90 (19)	3 (1)	0	5 (2)	0	0	NR	NR	NR	NR	NR	NR
Constipation	87 (19)	2 (<1)	0	45 (19)	0	0	65 (17)	1 (<1)	0	22 (11)	1 (1)	0
Abdominal pain	83 (18)	7 (1)	1 (<1)	60 (25)	10 (4)	0	105 (28)	13 (3)	NA	43 (22)	8 (4)	NA
Weight loss	81 (17)	5 (1)	0	14 (6)	0	0	51 (14)	7 (2)	NA	9 (5)	0	NA
Increase in ALT level	80 (17)	23 (5)	0	13 (5)	5 (2)	0	55 (15)	10 (3)	2 (1)	22 (11)	5 (3)	0
Mucosal inflammation†	65 (14)	8 (2)	0	5 (2)	1 (<1)	0	47 (13)	4 (1)	0	6 (3)	1 (1)	0
Pyrexia	64 (14)	0	0	24 (10)	1 (<1)	0	72 (9)	0	0	14 (7)	0	0
Upper abdominal pain	63 (13)	3 (1)	0	31 (13)	0	0	NR	NR	NR	NR	NR	NR
Cough	63 (13)	1 (<1)	0	26 (11)	0	0	40 (11)	1 (<1)	NA	14 (7)	0	NA
Peripheral oedema**	63 (13)	4 (1)	0	32 (14)	2 (1)	0	60 (16)	2 (1)	NA	24 (12)	0	NA
Stomatitis	63 (13)	8 (2)	0	5 (2)	0	0	NR	NR	NR	NR	NR	NR
Dyspnoea	58 (12)	15 (3)	0	24 (10)	1 (<1)	0	NR	NR	NR	NR	NR	NR
Rash	58 (12)	2 (<1)	0	14 (6)	1 (<1)	0	NR	NR	NR	NR	NR	NR
Ascites	57 (12)	17 (4)	1 (<1)	30 (13)	11 (5)	0	58 (16)	16 (4)	0	31 (16)	11 (6)	0
Dysgeusia	56 (12)	0	0	5 (2)	0	0	NR	NR	NR	NR	NR	NR
Hypoalbuminemia	55 (12)	2 (<1)	0	12 (5)	0	0	57 (15)	6 (2)	0	16 (8)	1 (1)	0
Headache	52 (11)	1 (<1)	0	16 (7)	1 (<1)	0	NR	NR	NR	NR	NR	NR
Insomnia	49 (10)	1 (<1)	0	17 (7)	0	0	NR	NR	NR	NR	NR	NR
Dizziness	48 (10)	2 (<1)	0	15 (6)	0	0	NR	NR	NR	NR	NR	NR
Dyspepsia	47 (10)	0	0	7 (3)	0	0	NR	NR	NR	NR	NR	NR
Anaemia	46 (10)	18 (4)	1 (<1)	19 (8)	12 (5)	0	58 (16)	16 (4)	2 (1)	22 (11)	0	NA
Back pain	46 (10)	5 (1)	0	24 (10)	1 (<1)	0	42 (11)	6 (2)	1 (<1)	17 (9)	2 (1)	0
Increase in serum bilirubin level	45 (10)	10 (2)	4 (1)	17 (7)	2 (1)	2 (1)	108 (29)	37 (10)	2 (1)	34 (19)	15 (8)	6 (3)

Decrease in platelet count	45 (10)	13 (3)	4 (1)	7 (3)	2 (1)	0	29 (10)	13 (3)	1 (<1)	5 (3)	0	0

* Listed are adverse events, regardless of causality, that were reported in at least 10% of patients in either group. Severity was graded according to National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0.
† Mucosal inflammation reported in CELESTIAL, whereas in RESORCE oral mucositis reported.
** Peripheral oedema reported in CELESTIAL, whereas in RESORCE limb oedema recorded.
Abbreviations: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; NA, not applicable; NR, not reported; PPES, palmar-plantar erythrodysaesthesia syndrome.
Source: Abou-Alfa et al., 2018, Bruix 2017, Exelixis, 2018.

A8. CS, Section B.3.12, page 83. The CS states *"No relevant studies are underway that are anticipated to provide additional evidence within the next 12 months to support the use of cabozantinib for the treatment of advanced HCC."* Please state whether there are any other ongoing or planned studies of cabozantinib or regorafenib relevant to this appraisal, irrespective of the date of expected results.

Company response

No, there are none in this post sorafenib setting.

A9. CS, Appendix E, Table 22 and Figure 6, pages 97 to 100. Subgroup analyses for CELESTIAL: The hazard ratio (HR) for overall survival (OS) is approximately 1 (i.e. no OS benefit for cabozantinib versus placebo) in the following subgroups: patients from Asia; patients with no extrahepatic spread, and patients with hepatitis C virus. Please comment on why this might be.

Company response

Subgroup analyses of progression-free survival suggested that cabozantinib had clinical activity across subgroups of patients with various aetiologic factors and across subgroups with other baseline characteristics. Subgroup analyses of overall survival were more variable, with broader confidence intervals. Hazard ratios in subgroups can be affected by statistical variability from evaluation of smaller populations or imbalances in prognostic factors or subsequent anticancer therapies, e.g., in the CELESTIAL trial only 25% of the patients were from Asia while the rest were from Europe, USA, Canada, Australia and New Zealand. In the regorafenib RESORCE trial 40% of the patient population were from Asia.

The uncertainty in the analyses of the subgroups is also highlighted by the variation between the subgroups of those located in Asia and those of Asian race, as patients in the geographical location of Asia had a HR for OS of 1.01 (0.68-1.48; 95% CI) while those of Asian race had a HR of 0.86 (0.63-1.19, 95% CI). This could reflect different

healthcare systems in Asia compared to other regions for example. To add further detail to this, the geographical region Asia included patients from Hong Kong, South Korea, Singapore and Taiwan (the regorafenib RESORCE trial included patients from China, Japan, South Korea, Singapore, and Taiwan).

At Ipsen's advisory board on the 31st of November 2021, clinical experts considered the lack of a statistically significant OS benefit for patients from Asia was considered to be an idiosyncrasy, related to small sample size, rather than a genuine differential effect of race on OS versus PFS. These conclusions were reiterated for the subgroups without EHS and those with HCV, as there was no clinical explanation for the observed differences between subgroups for OS.

A10. Priority. CS, Appendix E, Table 22 and Figure 6, pages 97 to 100. Subgroup analyses for CELESTIAL. For second-line patients, OS and PFS are very similar to those for the intention-to-treat (ITT) population used in the Bucher indirect treatment comparison (ITC). However, for third-line patients, the results are less favourable, and this is likely to be where cabozantinib would be used in the NHS. This may affect relative effectiveness and costs. Please comment.

Company response

Approximately 28% of patients in the cabozantinib arm in the CELESTIAL trial had two prior systemic anticancer regimens (third line) and thus the low patient numbers in the third line subgroup make it difficult to show a powered OS. The HR for OS was 0.9 (0.63-1.29, 95% CI) with the 95% confidence interval in third-line containing 1. Therefore, it cannot be concluded that cabozantinib increases OS relative to placebo. Yet it can be concluded that there is benefit in PFS with a hazard ratio of 0.58 (0.41-0.83, 95% CI).

There will naturally also be some uncertainty on whether the results in third line for cabozantinib in the CELESTIAL trial will be reflective in clinical practice in England and Wales, as the prior systemic anticancer therapies in the CELESTIAL trial (table 4) do not contain the combination of atezolizumab plus bevacizumab which has become the predominant first-line therapy for aHCC.

Table 4: CELESTIAL trial – prior systemic anticancer therapy.

Prior systemic anticancer therapy, n (%)	Cabozantinib (N = 470)	Placebo (N = 237)
Sorafenib	470 (100)	237 (100)
Regorafenib	6 (1)	2 (1)
Lenvatinib	0	1 (<1)
Tivantinib	1 (<1)	2 (1)
Ramucirumab	8 (2)	1 (<1)
Anti-PD-1/PD-L1	14 (3)	3 (1)
Cytotoxic chemotherapy	41 (9)	30 (13)
Doxorubicin	22 (5)	10 (4)
Investigational agent	60 (13)	20 (8)

Source: Abou-Alfa et al. 2018.

A key difference between the CELESTIAL and RESORCE trials is the inclusion of third-line patients in CELESTIAL whereas the RESORCE trial had none, and CELESTIAL has demonstrated a PFS benefit in third-line. Despite there being no evidence to support the use of regorafenib in third line, it is being used in practice in this line of therapy and funded by NHS England. Ipsen accepts the uncertain evidence for an OS benefit for cabozantinib in third line, but given the efficacy shown in PFS and the consideration that regorafenib is being prescribed at third line by NHS England and Wales without evidence, Ipsen would welcome a recommendation for use in third-line where regorafenib is already used.

Indirect comparisons

A11. Priority. CS, Section B.2, Table 6, page 30 and CELESTIAL trial publication (Abou-Alfa *et al.*, NEJM, 2018). Some patients in the RESORCE and CELESTIAL trials continued to receive their assigned treatment beyond disease progression. Patients in both trials may have also received subsequent anticancer therapies. Please comment on the extent to which these issues might confound the results of the comparison of OS outcomes from the ITCs.

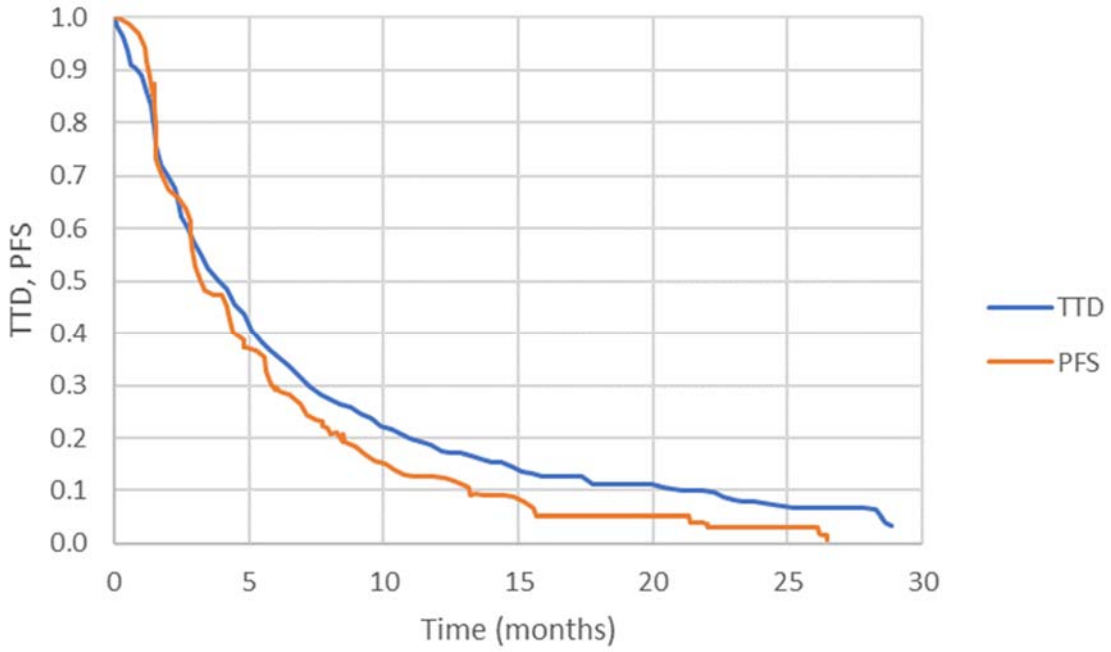
Company response

In the CELESTIAL trial, 26% of cabozantinib patients went on to have subsequent treatments; however, the number of patients that received third-line cabozantinib and received subsequent therapy is unknown. In the RESORCE trial, 20% of regorafenib

patients received subsequent therapy. Since the number of patients receiving subsequent treatment is relatively small and equivalent across both trials, the effect of subsequent treatment on the OS endpoint is expected to be limited.

Both the CELESTIAL and RESORCE trials allowed treatment beyond progression; however, this was more pronounced for regorafenib as shown in the comparison of the TTD and PFS KM curves for both trials (Figure 2 and 3). It is possible this may have introduced a bias towards an improvement for regorafenib in terms of its OS endpoint as evaluated in the RESORCE trial.

Figure 2: Comparison of the regorafenib treatment arm from RESORCE, PFS and TTD



Abbreviations: PFS, progression-free survival; TTD, time to treatment discontinuation
Note: patients on treatment on 29th February are considered censored
Source: Bruix et al. 2016; NICE TA555

Figure 3: Comparison of the cabozantinib treatment arm from CELESTIAL, PFS and TTD

Abbreviations: PFS, progression-free survival; TTD, time to treatment discontinuation
Source: CELESTIAL individual patient level data

A12. Priority. CS, Section B.3.10, page 59. The CS includes ITCs using the Bucher approach, anchored MAICs using constant and time-varying HRs and unanchored MAICs for PFS and OS. The results of the comparisons for OS are not fully consistent across all analyses. Please clarify which ITC analysis should be considered as the company's base case?

Company response

The Company performed an anchored MAIC, as well as a number of MAIC sensitivity analyses to address the issues with violation of the proportional hazards assumption and the ERG clarifications in this document. The Company has chosen the anchored MAIC based on the Weibull model approach as the base case because the basic underlying assumption (conditional constancy of relative effects) is easier to defend than the assumption of unanchored (constancy of absolute effects) and its conservative nature for cabozantinib, which is also the reason it will likely be the preferred scenario of the ERG and the NICE committee. The unanchored MAIC is also a relevant option to be considered given the violation of the PH assumption required

of the anchored MAIC. Further, the unanchored MAIC aligns with the findings of the ITC including real world evidence (RWE) for regorafenib in a population more closely resembling the cabozantinib population with respect to sorafenib tolerability, i.e., the RWE MAIC study by Casadei-Gardini et al (2021) which included patients irrespective of whether they tolerated sorafenib or not. Thus, the results of this published RWE MAIC study (see CS, Section B.3.10.4, Table 32) are more helpful in interpreting the positioning of cabozantinib in this submission within its marketing authorisation. To that end, the Company considers all MAICs presented to the ERG to be relevant options, reflecting the convergence of results demonstrating no meaningful difference in treatment effects between cabozantinib and regorafenib in a pure second line HCC population previously treated with sorafenib irrespective of tolerability.

A13. **Priority**. CS, Section B.3.10.2, page 60. Please provide the results from a Bucher ITC using the second-line population from both CELESTIAL and RESORCE (HRs provided in CS Appendix E, Figure 6).

Company response

The Bucher ITC results for cabozantinib versus regorafenib in the second-line population include a PFS HR of 0.93 (95% CI, 0.69-1.25) and an OS HR of 1.13 (95% CI, 0.83-1.53), suggesting no statistically significant difference between the two treatments in terms of PFS and OS. This analysis uses the RECIST 1.1 criteria for PFS as per clarification question A26 as well as the latest data cut for OS as per clarification question A25.

A14. **Priority**. CS, Section B.3.10.3, Table 27, page 67. Please provide details regarding the unweighted sample size (for both cabozantinib and placebo plus BSC arms) of the subpopulation of HCC patients who had prior treatment with sorafenib (i.e., pure second-line patients) in the CELESTIAL trial which was utilised in the MAIC analysis.

Company response

After removing subjects with missing values for the characteristics, the pure second line CELESTIAL population was reduced from 495 to 484 patients (326 in the cabozantinib arm and 158 in the placebo arm).

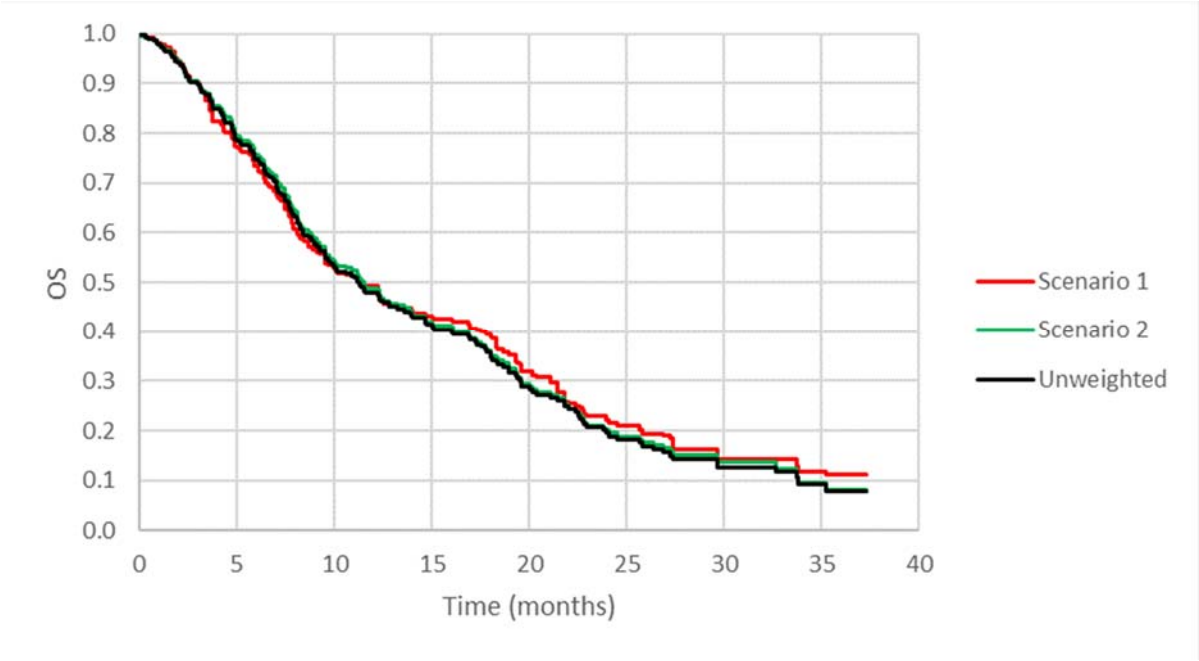
A15. Priority. CS, Section B.3.10.3, pages 65 to 78. For both the anchored and unanchored ITCs conducted for OS and PFS, please provide:

- A plot of unweighted and weighted Kaplan-Meier curves for CELESTIAL. Please plot the unweighted and weighted curves on the same figure;
- A plot of the empirical/unsmoothed and smoothed hazard function for the data used in the analysis. Please also plot the hazard function of the best fitting parametric model on top of the empirical and smoothed hazard;
- Where parametric survival models were fitted, please plot the fitted survival models together with the Kaplan-Meier curves;
- Specifically for the unanchored comparisons, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) estimates obtained from each of the parametric models fitted to the data for both OS and PFS.

Company response

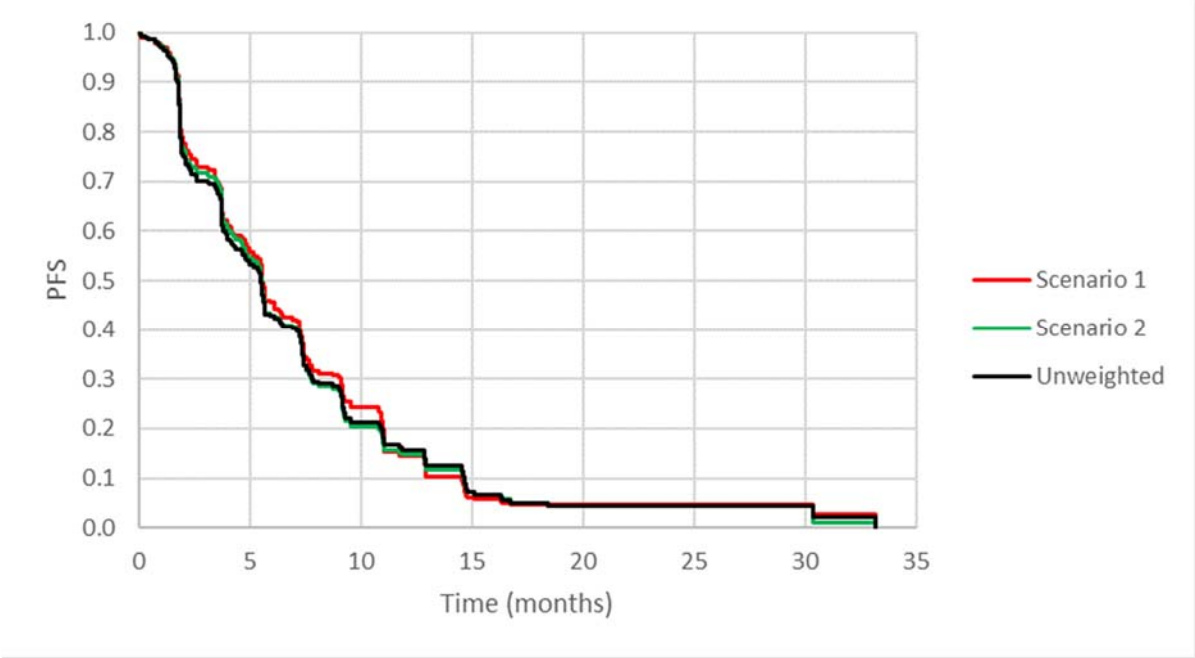
A comparison of both weighted KM scenarios from the MAIC and the unweighted pure second-line KM is shown in Figure 4 and Figure 5 for OS and PFS, respectively. As concluded in CS, Section B3.10.3, page 69, the scenarios are similar. The baseline characteristics selected for matching in scenario 1 were all potential effect modifiers identified by clinical expert opinion. This scenario was considered as the base case. Scenario 2 matches baseline characteristics identified using a stepwise AIC regression strategy. This scenario serves as a sensitivity analysis.

Figure 4: Weighted and unweighted cabozantinib OS KM



Abbreviations: KM, Kaplan-Meier; OS, overall survival

Figure 5: Weighted and unweighted cabozantinib PFS KM

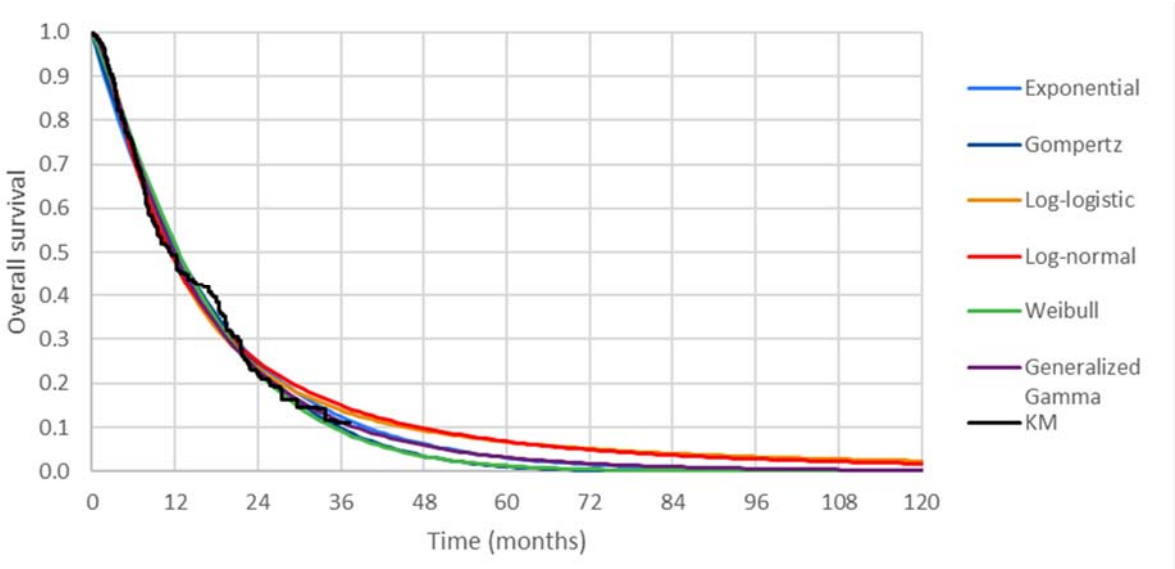


Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival

The parametric fits for the weighted cabozantinib OS data are shown in Figure 6, and the parametric fits for the regorafenib OS from the RESORCE trial are shown in Figure 7. The AIC and BIC estimates are shown in Table 5.

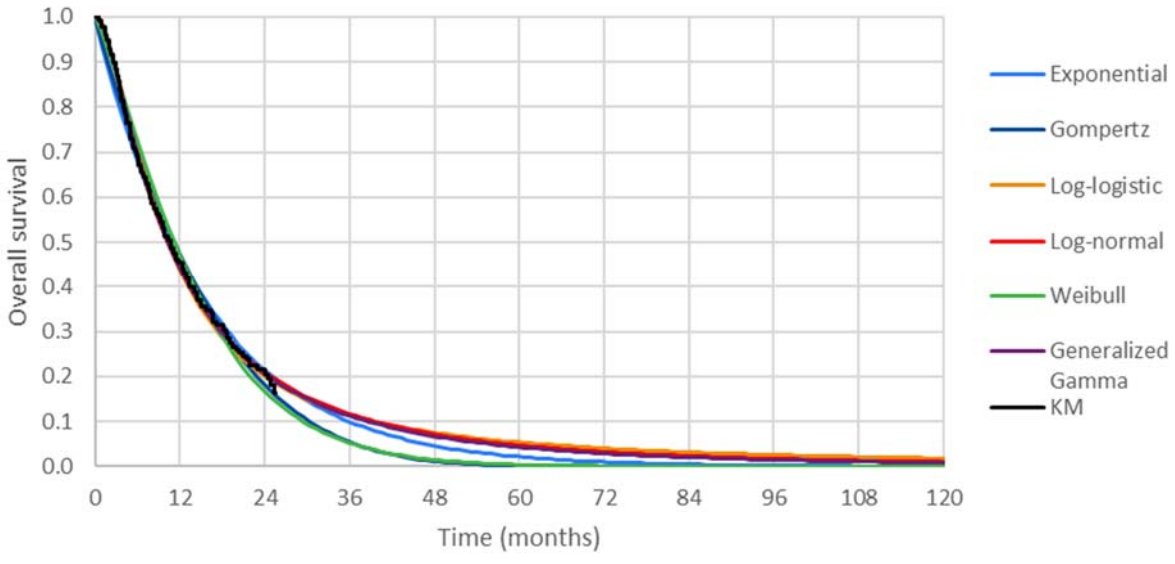
[CS Correction]: Please note a correction in the CS, Section B.3.10.3, page 76, that the generalised gamma is incorrectly labelled as the base case model for OS in the unanchored MAIC. The company would like to correct this error as it should be the log-logistic model. The CS, Section B.3.10.3, Figure 20 and reported mean and median OS are correct and correspond to the log-logistic model. The log-logistic distribution is selected to model the OS outcome as it appears to fit the weighted cabozantinib data better upon visual assessment.

Figure 6: Parametric fits for weighted cabozantinib OS



Abbreviations: KM, Kaplan-Meier; OS, overall survival

Figure 7: Parametric fits for regorafenib OS from the RESORCE trial



Abbreviations: KM, Kaplan-Meier; OS, overall survival

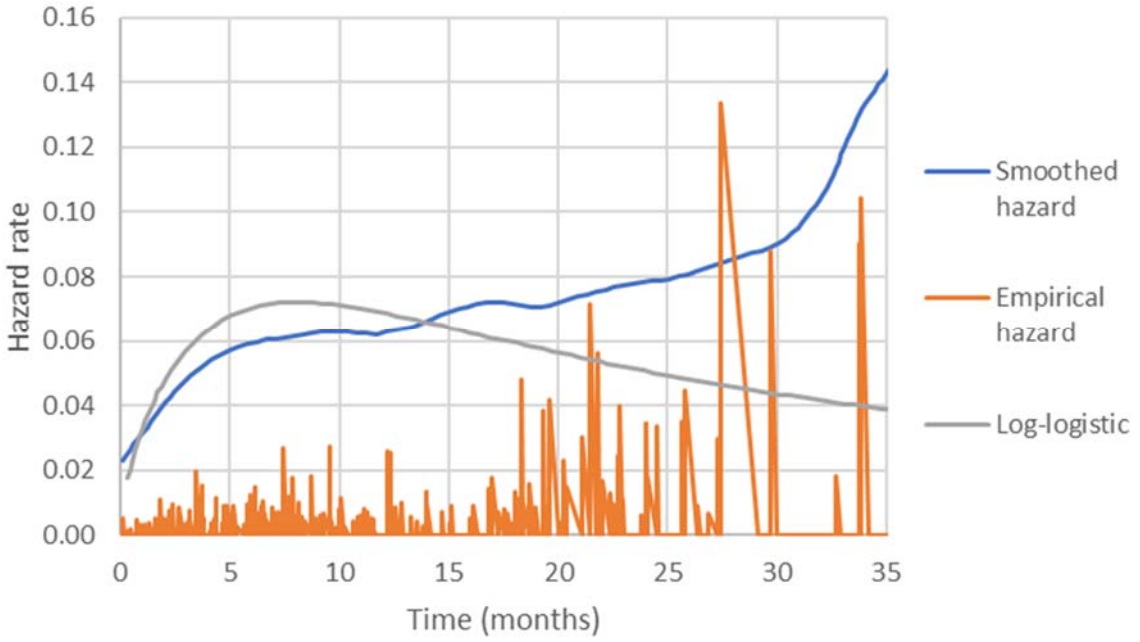
Table 5: AIC and BIC statistics for weighted cabozantinib and regorafenib OS parametric fits

	Weighted cabozantinib		Regorafenib	
Model	**AIC**	**BIC**	**AIC**	**BIC**
Exponential	1678.56	1682.34	1740.62	1744.56
Weibull	1672.09	1679.67	1727.96	1735.84
Gompertz	1678.39	1685.96	1739.24	1747.11
log-logistic	1668.20	1675.78	1716.81	1724.68
log-normal	1675.18	1682.75	1712.17	1720.05
Generalised gamma	1668.37	1679.74	1714.10	1725.92

Abbreviations: AIC, akaike information criterion; BIC, bayesian information criterion.; OS, overall survival

The hazard of the base case OS model and the empirical and smoothed hazard of the OS KM is shown in Figure 8 and Figure 9 for cabozantinib and regorafenib, respectively. The smoothed hazard was generated from the flexsurv package in R which utilises the 'muhaz' package in hazard calculations. The 'muhaz' package does not take into account the weights from the MAIC and no alternative was readily available. Therefore, the smoothed hazard presented for the weighted cabozantinib OS KM has a limitation.

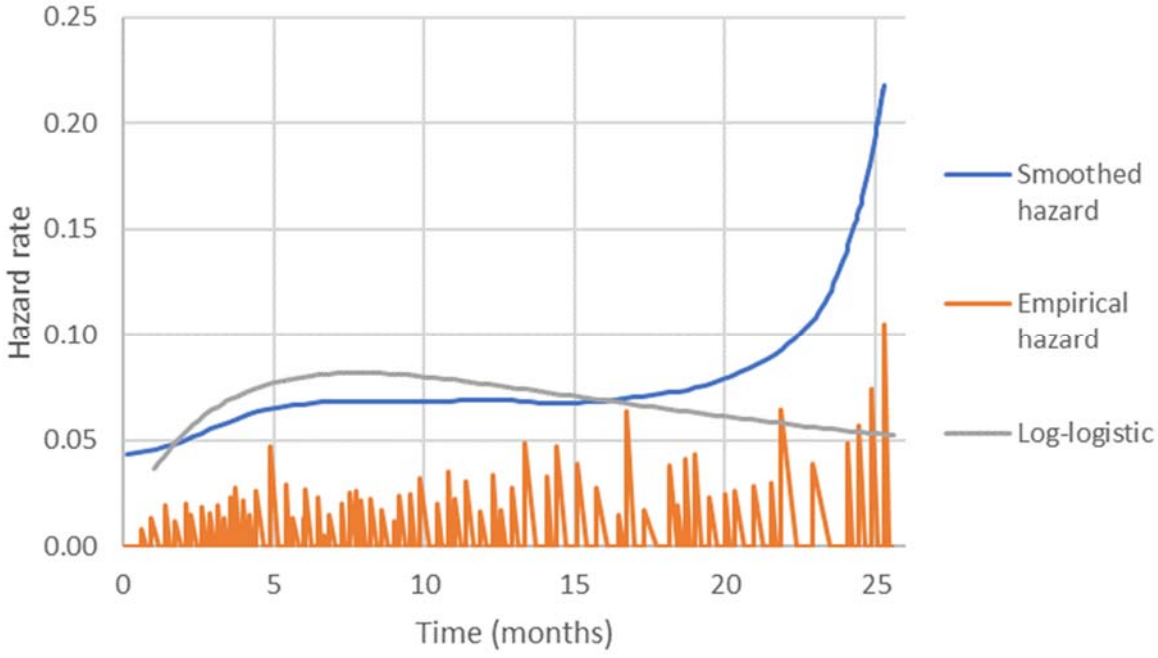
Figure 8: Weighted cabozantinib OS hazard rate



Abbreviations: KM, Kaplan-Meier; OS, overall survival

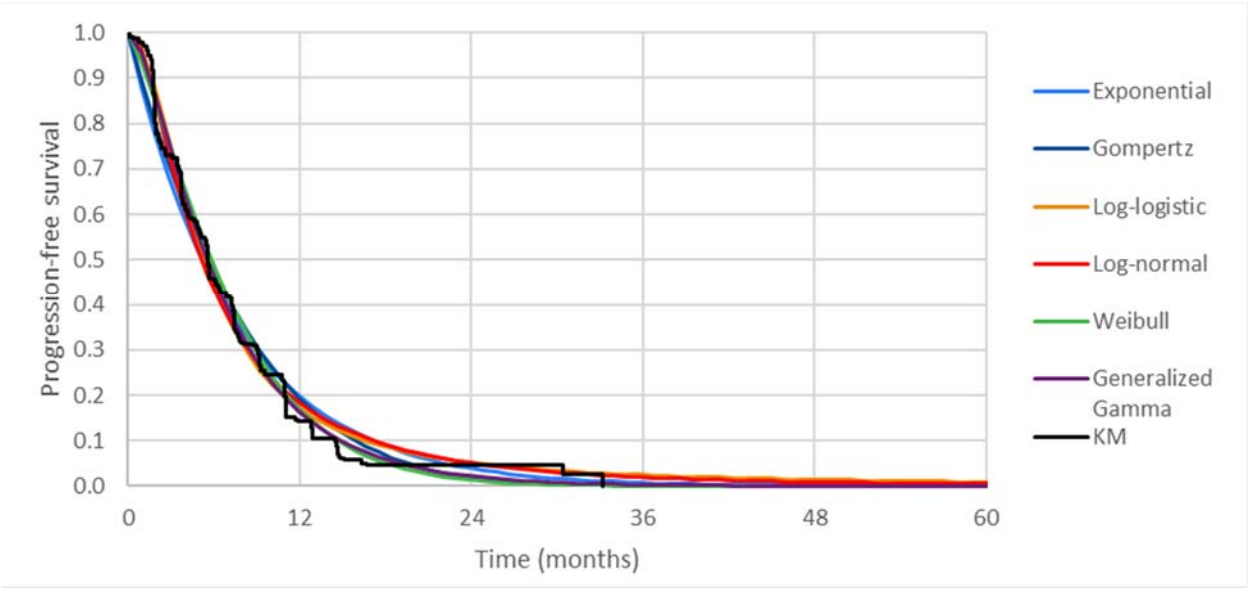
Figure 9: Regorafenib OS hazard rate from the RESORCE trial

Abbreviations: KM, Kaplan-Meier; OS, overall survival



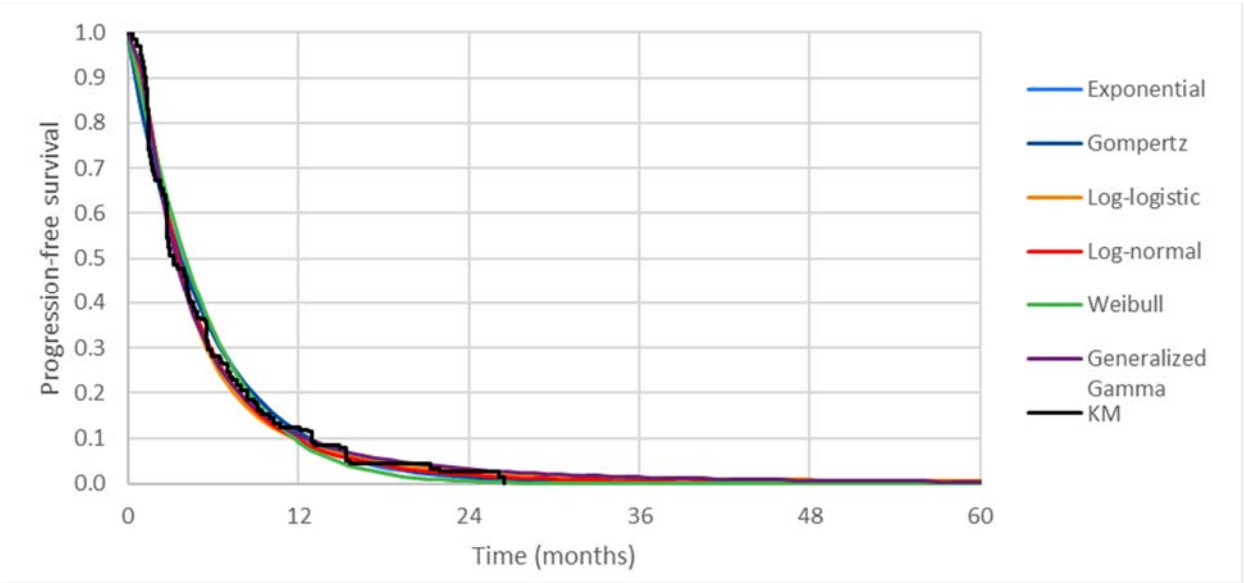
The parametric fits for the weighted cabozantinib PFS data are shown in Figure 10 and the parametric fits for the regorafenib PFS from the RESORCE trial are shown in Figure 11. The AIC and BIC statistics are shown in Table 6. The generalised gamma is selected as the base case model due to the better statistical fit.

Figure 10: Parametric fits for weighted cabozantinib PFS



Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival

Figure 11: Parametric fits for regorafenib PFS from the RESORCE trial



Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival

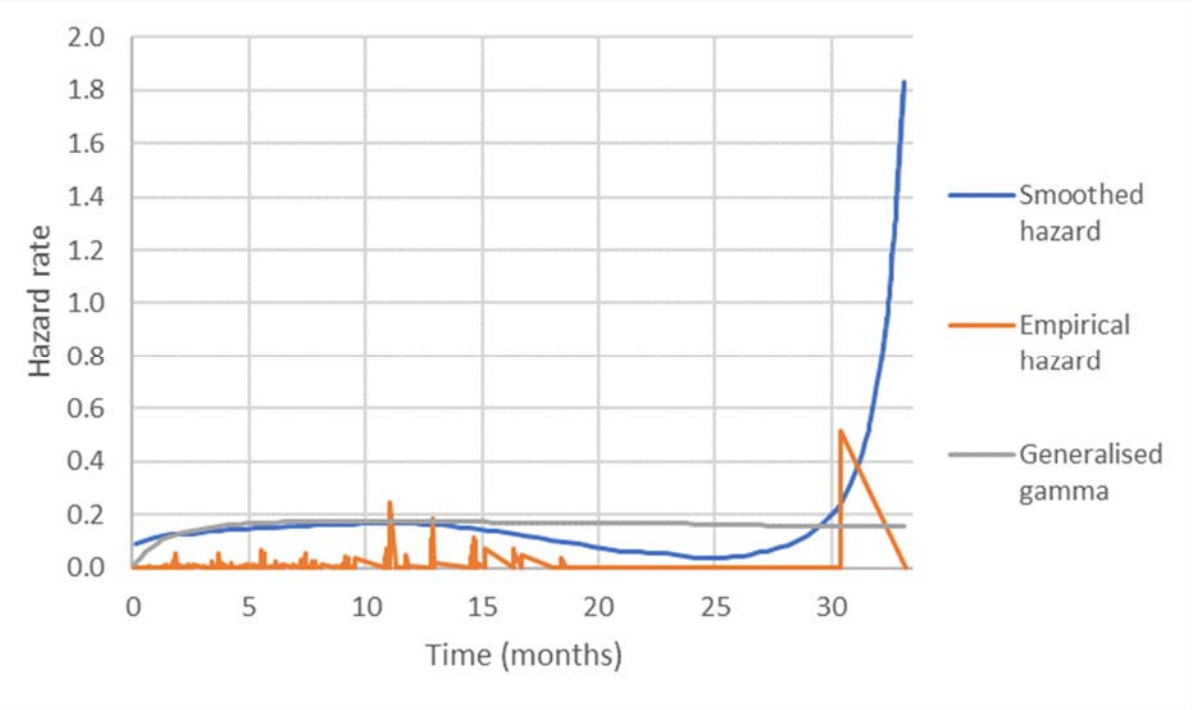
Table 6: AIC and BIC statistics for weighted cabozantinib and regorafenib PFS parametric fits

	Weighted cabozantinib		Regorafenib	
Model	**AIC**	**BIC**	**AIC**	**BIC**
Exponential	1480.30	1484.09	1641.66	1645.60
Weibull	1457.16	1464.73	1634.92	1642.79
Gompertz	1476.18	1483.75	1643.38	1651.26
log-logistic	1453.83	1461.41	1590.28	1598.15
log-normal	1467.01	1474.58	1577.40	1585.27
Generalised gamma	1450.61	1461.97	1575.13	1586.94

Abbreviations: AIC, akaike information criterion; BIC, bayesian information criterion.; PFS, progression-free survival

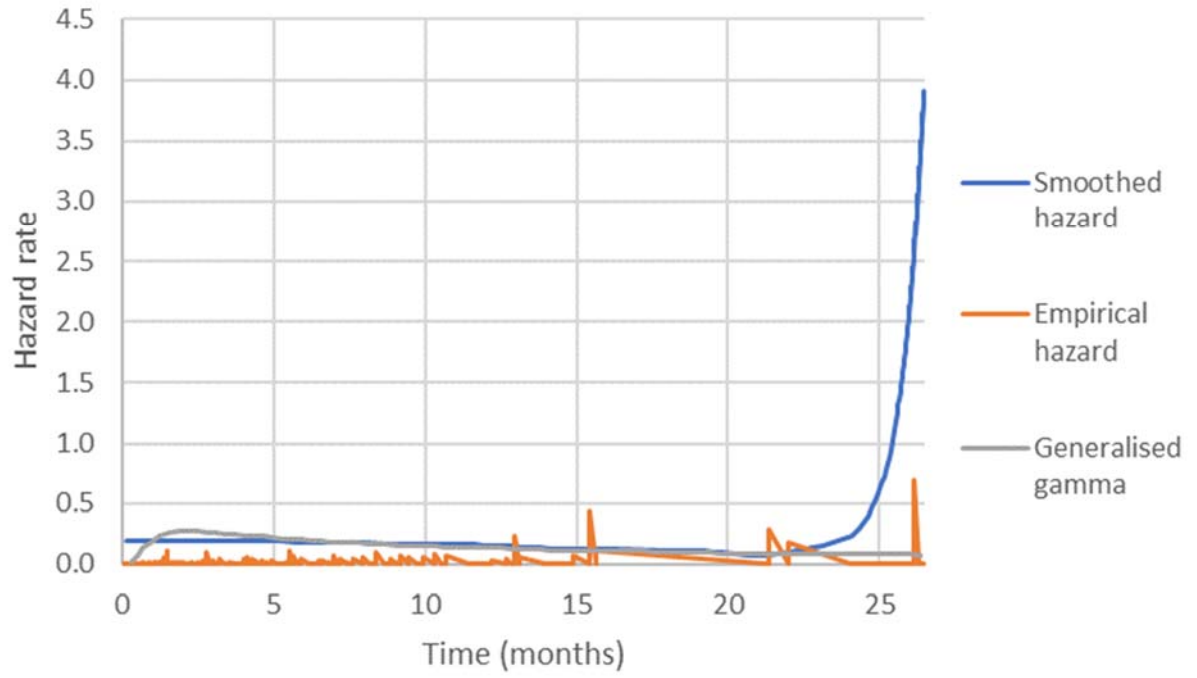
The hazard of the base case model and the empirical and smoothed hazard of the KM is shown in Figure 12 and Figure for cabozantinib and regorafenib respectively. Similar to OS, the smoothed hazard of the weighted cabozantinib PFS KM has a limitation due to lack of utilisation of weights in the 'muhaz' package.

Figure 12: Weighted cabozantinib PFS hazard rate



Abbreviations: PFS, progression-free survival

Figure 13: Regorafenib PFS hazard rate from the RESORCE trial



Abbreviations: PFS, progression-free survival

A16. Priority. CS, Section B.3.10.2, Table 26, page 65. Should the odds ratios (ORs) presented in this table represent a comparison of regorafenib versus cabozantinib (i.e., is there a labelling error)?

Company response

The table is labelled correctly to show the OR of cabozantinib versus regorafenib. A breakdown of the calculation is shown in Table 7.

Table 7: AE OR calculations used in the Bucher analysis (ITT)

AE	Cabozantinib versus placebo OR*	Regorafenib versus placebo OR*	Cabozantinib versus regorafenib OR
Hypertension	(398/69)/(235/2) = 0.05	(325/49)/(187/6) = 0.21	0.05/0.21 = 0.2
Elevated aspartate aminotransferase	(431/36)/(226/11) = 0.58	(355/19)/(183/10) = 1.02	0.58/1.02 = 0.6
Fatigue	(428/39)/(231/6) = 0.29	(350/24)/(190/3) = 0.23	0.29/0.23 = 1.2

***Note**: Calculated as AE odds of intervention / odds of placebo
Abbreviations: AE, adverse event; OR, odd ratio

A17. CS, Section B.3.10.3, Table 27, page 67. Please confirm if the footnote regarding the approximate effective sample size should be assigned to the sample sizes given in the header of Table 27.

Company response

The approximate effective sample sizes in CS, Section B.3.10.3, Table 27, page 67, refer to both of the MAIC baseline characteristics scenarios for cabozantinib. Therefore, the approximate effective sample size for cabozantinib in scenario 1 and scenario 2 is N = 187.27 and N = 303.24, respectively. The sample size for regorafenib is as reported, N = 374.

A18. Priority. CS, Section B.3.10.3, Table 28, page 68. Please provide details of the classification of each of the prognostic factors and treatment effect modifiers included in the matching process and provide details on how this classification was determined for each factor.

Company response

Table 8 below provides details regarding the classification of each of the included prognostic factors and treatment effect modifiers and the associated rationale.

Proportions and means were published for RESORCE for the following characteristics and were available for CELESTIAL: age group, gender, geographical region, ECOG performance status, Child-Pugh class, duration of prior sorafenib treatment, extrahepatic disease, macrovascular invasion, extrahepatic disease and/or macrovascular invasion, aetiology of disease, AFP level and race. Patients recruited to the RESORCE trial had increased tolerability to sorafenib; however, due to lack of reported data, this variable could not be accounted for. Considering the limited data, duration of prior sorafenib was considered as a proxy for sorafenib tolerability.

All the aforementioned characteristics were presented as dichotomous variables for RESORCE (Bruix et al., 2017), except for duration of prior sorafenib treatment, aetiology of disease and race. Duration of prior sorafenib treatment is a continuous variable, aetiology of disease is a categorical variable with 6 categories (Hepatitis B, alcohol use, Hepatitis C, unknown, non-alcoholic steatohepatitis and other), and race is a categorical variable with 4 categories (white, Asian, black and other/not reported).

Duration of prior sorafenib treatment has been reported as a mean (Finn et al., 2018), and all the dichotomous and categorical variables have been reported as percentages. Aetiology and race have been dichotomised into multiple characteristics (one for each category).

The following measures were taken to remove characteristics suspected to introduce noise into the matching processor to be strongly correlated with other baseline covariates. For aetiology and race, categories including under 10% of subjects in each trial (such as non-alcoholic steatohepatitis aetiology and black race) or representing 'other' or 'not reported' were not matched. Asian race is very strongly correlated with Asia geographical region (e.g., all the patients enrolled in the CELESTIAL trial in Asia are of Asian race); it is therefore not matched. Similarly, 'extrahepatic disease and/or macrovascular invasion' is not matched as it is evidently very strongly correlated with each of its individual components. Matching the aforementioned characteristics would likely result in a loss of statistical power/efficiency and overmatched/overfitted data (if a covariate is already balanced across the two trials, except for random noise, matching it will just introduce additional noise into the system).

Table 8: Classification of prognostic factors and treatment effect modifiers used in matching

Characteristic	Classification	Rationale for classification
Age	< 65 years ≥ 65 years	To reflect average age in RESORCE. This was categorised to minimise impact on effective sample size
Sex	Female Male	Binary variable
Region	Asia Other	To reflect reporting of RESORCE trial region baseline characteristic
ECOG status	ECOG 0 ECOG 1 or 2	Binary variable. ECOG 1 and ECOG 2 combined due to low ECOG 2 numbers
Child-Pugh score	A B or unknown	Binary variable
Mean duration of prior sorafenib	Continuous variable	-
Disease status: Extrahepatic disease	Present Absent	Binary variable
Disease status: Macrovascular invasion	Present Absent	Binary variable

Characteristic	Classification	Rationale for classification
	Unknown	
Aetiology of disease: Hepatitis B	Present Absent Unknown	Binary variable
Aetiology of disease: Alcohol use	Present Absent Unknown	Binary variable
Aetiology of disease: Hepatitis C	Present Absent Unknown	Binary variable
Alpha fetoprotein level	≥ 400 ng/ml < 400 ng/ml	To reflect reporting of RESORCE trial alpha fetoprotein level baseline characteristic. This is a diagnostic threshold for HCC
Race: White	Yes No	Binary variable

Abbreviations: ECOG, Eastern Cooperative Oncology Group

A19. CS, Section B.3.10.3, page 70. Please present the results of a MAIC for "Any G3/4 AEs" for regorafenib and cabozantinib using data from CELESTIAL and RESORCE.

Company response

Grade 3/4 treatment-emergent drug-related adverse events affecting >5% of the subjects on any arm of either CELESTIAL or RESORCE were compared between cabozantinib versus regorafenib in the MAIC. These adverse events were: hypertension, aspartate aminotransferase increase (called increased AST in RESORCE), fatigue, diarrhoea, palmar-plantar erythrodysesthesia syndrome (called hand-foot skin reaction in RESORCE) and elevated bilirubin. A MAIC was performed for the individual grade 3 / 4 AEs and not any grade 3 /4 AEs for two reasons: 1) the incidence of any grade 3 / 4 AEs in the ITT population was almost identical between the two treatment arms and a MAIC would similarly be expected to yield no difference; and 2) the MAIC AE results were intended to serve as input into a cost-utility analysis allowing any meaningful differences to be reflected in terms of disutility and treatment costs. A side-by-side comparison of the adverse events in the ITT population as reported in the pivotal trial publications (Table 3) demonstrates similarity in terms of

incidence of adverse events; similar safety profiles between the two treatment is further supported by clinical experts.

It is also important to note that despite similar safety findings between the two trials, there may have been underreporting of certain adverse events and especially serious adverse events in HCC in the RESORCE trial since patients that discontinued prior sorafenib due to sorafenib-related toxicity were excluded from the study and sorafenib belongs to the same pharmacological class. The impact on the reported safety profile for regorafenib in HCC is therefore unknown. A warning was added to section 4.4 of regorafenib's SmPC to reflect this limitation.

A20. Priority. CS, Section B.3.10.3, page 75; CS Appendix I, page 111. The CS states that an anchored comparison was conducted *"by generating time-varying hazard ratios from hazard profiles of fitted parametric models to the weighted CELESTIAL and RESORCE data"* and *"the fitted hazard functions of each treatment were used to generate HRs over time."* Please provide more detail around the time-varying HR ITC approach, specifically:

- The methodology adopted to generate the HR after fitting independent parametric curves to the data;
- How the weights from the matching have been incorporated into the indirect comparison;
- The statistical programming code used to estimate the time-varying HRs.

Company response

Independent parametric models were fitted to the weighted cabozantinib data, weighted CELESTIAL placebo data, regorafenib data and RESORCE placebo data in order to generate the hazard for each treatment arm. The time-varying cabozantinib versus CELESTIAL placebo HR was generated by dividing the hazard of the cabozantinib parametric model by the hazard of the CELESTIAL placebo parametric model at each timepoint. The regorafenib versus RESORCE placebo time-varying HR was generated in the same way. The time-varying HR of cabozantinib versus regorafenib was generated by calculating the ratio of the cabozantinib versus

CELESTIAL placebo HR with the regorafenib versus RESORCE placebo time-varying HR.

The population matching was incorporated into the analysis through the use of the weights in the parametric model fitting stage for cabozantinib and CELESTIAL placebo.

The R code used to generate this analysis is shown in Table 9.

Table 9: R code for time-varying hazard ratio analysis

Code

```
# Package setup
library(tidyverse)
library(flexsurv)

# Distribution inputs for flexsurvreg function
dists=c('exp', 'weibull', 'lnorm', 'llogis', 'gompertz', 'gengamma')

# Initializing empty list
mods=NULL

# Models on arm A of the population-adjusted AB data
mods[[1]]=lapply(dists, function(x) {
mod=AB_IPD %>% mutate(wt) %>% filter(trt=='A') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., weights=wt, dist=x)
})

# Models on arm B of the population-adjusted AB data
mods[[2]]=lapply(dists, function(x) {
mod=AB_IPD %>% mutate(wt) %>% filter(trt=='B') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., weights=wt, dist=x)
})

# Models on arm A of the digitized AC pseudo-data
mods[[5]]=lapply(dists, function(x) {
mod=AC_pseudo %>% filter(trt=='A') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., dist=x)
})

# Models on arm C of the digitized AC pseudo-data
mods[[6]]=lapply(dists, function(x) {
mod=AC_pseudo %>% filter(trt=='C') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., dist=x)
})

# Time horizon and timepoints for hazard extrapolations
horizon=15
times=seq(from=0, to=horizon, length=1000)

# Deriving adjusted log-hazard ratio from trial AB
# Log-hazard of arm A
log_haz_A_AB_MAIC=AB_IPD %>% mutate(wt) %>% filter(trt=='A') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., weights=wt, dist='weibull') %>%
```

Code

```
summary(t=times, type='hazard', tidy=T) %>%
mutate(est=log(est), lcl=log(lcl), ucl=log(ucl)) %>%
filter_all(all_vars(!is.infinite(.))) %>%
mutate(var=((ucl - lcl) / (qnorm(0.975) - qnorm(0.025)))^2) %>%
select(-c(lcl, ucl)) %>%
mutate(Trial='A_AB_MAIC') %>% as_tibble()

# Log-hazard of arm B
log_haz_B_AB_MAIC=AB_IPD %>% mutate(wt) %>% filter(trt=='B') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., weights=wt, dist='weibull') %>%
summary(t=times, type='hazard', tidy=T) %>%
mutate(est=log(est), lcl=log(lcl), ucl=log(ucl)) %>%
filter_all(all_vars(!is.infinite(.))) %>%
mutate(var=((ucl - lcl) / (qnorm(0.975) - qnorm(0.025)))^2) %>%
select(-c(lcl, ucl)) %>%
mutate(Trial='B_AB_MAIC') %>% as_tibble()

# Log-hazard ratio
log_hazard_ratio_AB_MAIC=rbind(log_haz_A_AB_MAIC, log_haz_B_AB_MAIC) %>%
pivot_wider(names_from=Trial, values_from=c(est, var)) %>%
mutate(logHR_AB_MAIC=est_B_AB_MAIC - est_A_AB_MAIC,
logHR_AB_MAIC_var=var_A_AB_MAIC + var_B_AB_MAIC) %>%
select(-c(2:5)) %>% mutate(Comparison='AB_AC') %>%
rename(logHR=logHR_AB_MAIC, logHR_var=logHR_AB_MAIC_var)

# Deriving log-hazard ratio from trial AC
# Log-hazard of arm A
log_haz_A_AC=AC_pseudo %>% filter(trt=='A') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., dist='weibull') %>%
summary(t=times, type='hazard', tidy=T) %>%
mutate(est=log(est), lcl=log(lcl), ucl=log(ucl)) %>%
filter_all(all_vars(!is.infinite(.))) %>%
mutate(var=((ucl - lcl) / (qnorm(0.975) - qnorm(0.025)))^2) %>%
select(-c(lcl, ucl)) %>%
mutate(Trial='A_AC') %>% as_tibble()

# Log-hazard of arm C
log_haz_C_AC=AC_pseudo %>% filter(trt=='C') %>%
flexsurvreg(Surv(Time, Event) ~ 1, data=., dist='weibull') %>%
summary(t=times, type='hazard', tidy=T) %>%
mutate(est=log(est), lcl=log(lcl), ucl=log(ucl)) %>%
filter_all(all_vars(!is.infinite(.))) %>%
mutate(var=((ucl - lcl) / (qnorm(0.975) - qnorm(0.025)))^2) %>%
select(-c(lcl, ucl)) %>%
mutate(Trial='C_AC') %>% as_tibble()

# Log-hazard ratio
log_hazard_ratio_AC=rbind(log_haz_A_AC, log_haz_C_AC) %>%
pivot_wider(names_from=Trial, values_from=c(est, var)) %>%
mutate(logHR_AC=est_C_AC - est_A_AC,
logHR_AC_var=var_A_AC + var_C_AC) %>%
select(-c(2:5)) %>% mutate(Comparison='AC_AC') %>%
rename(logHR=logHR_AC, logHR_var=logHR_AC_var)

# Calculating adjusted time-varying hazard ratios of C vs B
HR_BC_MAIC=rbind(log_hazard_ratio_AB_MAIC, log_hazard_ratio_AC) %>%
pivot_wider(names_from=Comparison, values_from=c(logHR, logHR_var)) %>%
mutate(logHR_BC=logHR_AC_AC - logHR_AB_AC,
```

Code
logHR_BC_var=logHR_var_AC_AC + logHR_var_AB_AC, logHR_BC_lo=logHR_BC + qnorm(0.025) * sqrt(logHR_BC_var), logHR_BC_hi=logHR_BC + qnorm(0.975) * sqrt(logHR_BC_var)) %>% select(time, logHR_BC, logHR_BC_lo, logHR_BC_hi) %>% mutate(Comparison='MAIC')

A21. CS, Section B.3.10.3 Table 31, page 75 and Figures 17-20, pages 76-78. Please confirm if the weighted results from the anchored and unanchored ITCs utilise weights from Scenario 1 or Scenario 2.

Company response

The anchored and unanchored ITCs utilise weights from scenario 1.

A22. CS, Section B.3.10.3, page 77. Please provide the coefficients estimated from each of the parametric survival models fitted to the data and provide details of which software package was utilised to produce the survival curves.

Company response

The 'survey' and 'flexsurv' packages in R were used to generate survival curves. Custom functions derived from these packages were used to fit generalised gamma and gompertz models and obtain AIC/BIC values for weighted parametric models. The statistical programming code for the custom functions is provided in Table 10. Table 11 shows which packages were used for each model.

Table 10: Statistical packages used for survival analysis

Data & Model	Package & function
RESORCE, **Exponential** **Weibull** **Lognormal** **Log-logistic**	Survey survreg()
RESORCE **Generalised gamma** **Gompertz**	Flexsurv flexsurvreg.edit()
Weighted CELESTIAL **Exponential** **Weibull** **Lognormal** **Log-logistic**	Survey svysurvreg.survey.design()
Weighted CELESTIAL **Generalised gamma**	Flexsurv svyflexsurvreg.survey.design()

Data & Model	Package & function
Gompertz	

Table 11: R code for custom functions used in survival analysis

Code

```
## sinh(log(y))
logh <- function(x) { 0.5 * (x - 1/x) }

buildTransformer <- function(inits, nbpars, dlist) {
  par.transform <-
    lapply(seq_len(nbpars),
           function(ind) {
             xform  <- dlist$inv.transforms[[ind]]
             function(pars) {
               xform(pars[[ind]])
             }
           })
  names(par.transform) <- names(inits)[seq_len(nbpars)]
  function(pars) {
    lapply(par.transform,
           function(item, par) { item(par) },
           pars)
  }
}

buildAuxParms <- function(aux, dlist) {
  aux.transform <- list()
  for (ind in seq_along(aux)) {
    name <- names(aux)[[ind]]
    if (!(name %in% dlist$pars)) {
      aux.transform[[name]] <- aux[[ind]]
    }
  }
  aux.transform
}

logLikFactory <- function(Y, X=0, weights, bhazard, dlist,
                          inits, dfns, aux, mx, fixedpars=NULL) {
  pars   <- inits
  npars  <- length(pars)
  nbpars <- length(dlist$pars)
  insert.locations <- setdiff(seq_len(npars),
                              fixedpars)

  ## which are the subjects with events
  event <- Y[,"status"] == 1
  event.times <- Y[event, "time1"]
  left.censor <- Y[!event, "time2"]
  right.censor <- Y[!event, "time1"]

  event.weights <- weights[event]
  no.event.weights <- weights[!event]

  par.transform <- buildTransformer(inits, nbpars, dlist)
```

Code

```
  aux.pars <- buildAuxParms(aux, dlist)

  default.offset <- rep.int(0, length(event.times))
  do.hazard <- any(bhazard > 0)

  loglik <- rep.int(0, nrow(Y))
  ## the ... here is to work around optim
  function(optpars, ...) {

    pars[insert.locations] <- optpars
    raw.pars <- pars
    pars <- as.list(pars)

    pars.event <- pars.nevent <- pars
    if (npars > nbpars) {
      beta <- raw.pars[(nbpars+1):npars]
      for (i in dlist$pars){
        pars[[i]] <- pars[[i]] +
          X[,mx[[i]],drop=FALSE] %*% beta[mx[[i]]]
        pars.event[[i]] <- pars[[i]][event]
        pars.nevent[[i]] <- pars[[i]][!event]
      }
    }

    fnargs <- c(par.transform(pars),
                aux.pars)
    fnargs.event <- c(par.transform(pars.event),
                      aux.pars)
    fnargs.nevent <- c(par.transform(pars.nevent),
                       aux.pars)

    ## Generic survival model likelihood contributions
    ## Observed deaths
    dargs <- fnargs.event
    dargs$x <- event.times
    dargs$log <- TRUE
    logdens <- do.call(dfns$d, dargs)

    ## Left censoring times (upper bound for event time)
    if (any(!event)){
      pmaxargs <- fnargs.nevent
      pmaxargs$q <- left.censor # Inf if right-censored, giving pmax=1
      pmax <- do.call(dfns$p, pmaxargs)
      pmax[pmaxargs$q==Inf] <- 1  # in case user-defined function doesn't already do this

      ## Right censoring times (lower bound for event time)
      pargs <- fnargs.nevent
      pargs$q <- right.censor
      pmin <- do.call(dfns$p, pargs)
    }

    ## Left-truncation
    targs   <- fnargs
    targs$q <- Y[,"start"]
    pobs <- 1 - do.call(dfns$p, targs) # prob of being observed = 1 unless left-truncated

    ## Hazard offset for relative survival models
    if (do.hazard){
```

Code

```
      pargs   <- fnargs.event
      pargs$q <- event.times
      pminb   <- do.call(dfns$p, pargs)
      loghaz  <- logdens - log(1 - pminb)
      offseti <- log(1 + bhazard[event] / exp(loghaz)*weights[event])
    } else {
      offseti <- default.offset
    }
    ## Express as vector of individual likelihood contributions
    loglik[event] <- (logdens*event.weights) + offseti
    if (any(!event))
      loglik[!event] <- (log(pmax - pmin)*no.event.weights)
    loglik <- loglik - log(pobs)*weights

    ret <- -sum(loglik)
    attr(ret, "indiv") <- loglik
    ret
  }
}

minusloglik.flexsurv <- function(optpars, Y, X=0, weights, bhazard,
                                 dlist, inits,
                                 dfns, aux, mx, fixedpars=NULL) {
  logLikFactory(Y, X, weights, bhazard, dlist, inits,
                dfns, aux, mx, fixedpars=fixedpars)(optpars)
}

check.dlist <- function(dlist){
  ## put tests in testthat
  if (is.null(dlist$name)) stop("\"name\" element of custom distribution list not given")
  if (!is.character(dlist$name)) stop("\"name\" element of custom distribution list should be a string")
  if (is.null(dlist$pars)) stop("parameter names \"pars\" not given in custom distribution list")
  if (!is.character(dlist$pars)) stop("parameter names \"pars\" should be a character vector")
  npars <- length(dlist$pars)
  if (is.null(dlist$location)) {
    warning("location parameter not given, assuming it is the first one")
    dlist$location <- dlist$pars[1]
  }
  if (!(dlist$location %in% dlist$pars)) {
    stop(sprintf("location parameter \"%s\" not in list of parameters", dlist$location))
  }
  if (is.null(dlist$transforms)) stop("transforms not given in custom distribution list")
  if (is.null(dlist$inv.transforms)) stop("inverse transforms not given in custom distribution list")
  if (!is.list(dlist$transforms)) stop("\"transforms\" must be a list of functions")
  if (!is.list(dlist$inv.transforms)) stop("\"inv.transforms\" must be a list of functions")
  if (!all(sapply(dlist$transforms, is.function))) stop("some of \"transforms\" are not functions")
  if (!all(sapply(dlist$inv.transforms, is.function))) stop("some of \"inv.transforms\" are not functions")
  if (length(dlist$transforms) != npars) stop("transforms vector of length ",length(dlist$transforms),",
parameter names of length ",npars)
  if (length(dlist$inv.transforms) != npars) stop("inverse transforms vector of length
",length(dlist$inv.transforms),", parameter names of length ",npars) #
  for (i in 1:npars){
    if (is.character(dlist$transforms[[i]])) dlist$transforms[[i]] <- get(dlist$transforms[[i]])
    if (is.character(dlist$inv.transforms[[i]])) dlist$inv.transforms[[i]] <- get(dlist$inv.transforms[[i]])
    if (!is.function(dlist$transforms[[i]])) stop("Transformation function for parameter ", i, " not
defined")
    if (!is.function(dlist$inv.transforms[[i]])) stop("Inverse transformation function for parameter ", i, "
not defined")
```

Code

```
  }
  if (!is.null(dlist$inits) && !is.function(dlist$inits)) stop("\"inits\" element of custom distribution list
must be a function")
  dlist
}
## Return formula for linear model on parameter called "par"
## Parameters should not have the same name as anything that might
## appear as a function in a formula (such as "I", "strata", or
## "factor").  If any parameters of the distribution being used are
## named like this, then such model functions will be interpreted as
## parameters and will not work

check.formula <- function(formula, dlist){
  if (!inherits(formula,"formula")) stop("\"formula\" must be a formula object")
  if (!("strata" %in% dlist$pars)){
    labs <- attr(terms(formula), "term.labels")
    strat <- grep("strata\\((.+)\\)",labs)
    if (any(strat)){
      cov <- gsub("strata\\((.+)\\)","\\1",labs[strat[1]])
      warning("Ignoring \"strata\" function: interpreting \"",cov, "\" as a covariate on \"", dlist$location,
"\"")
    }
  }
}

ancpar.formula <- function(formula, par){
  labs <- attr(terms(formula), "term.labels")
  pattern <- paste0(par,"\\((.+)\\)")
  labs <- grep(pattern,labs,value=TRUE)
  if (length(labs)==0) return(NULL)
  labs <- gsub(pattern, "\\1", labs)
  f2 <- reformulate(labs)
  environment(f2) <- environment(formula)
  f2
}

## Omit formula terms containing ancillary parameters, leaving only
## the formula for the location parameter

get.locform <- function(formula, ancnames){
  labs <- attr(terms(formula), "term.labels")
  dropx <- unlist(lapply(ancnames, function(x){grep(paste0(x,"\\((.+)\\)"),labs)}))
  formula(terms(formula)[c(0,setdiff(seq_along(labs),dropx))])
}

## Concatenate location formula (that includes Surv response term)
## with list of ancillary formulae, giving a merged formula to obtain
## the model frame

concat.formulae <- function(formula,forms){
  covnames <- unlist(lapply(forms, function(x)attr(terms(x),"term.labels")))
  covform <- if(length(covnames)==0) "1" else paste(covnames, collapse=" + ")
  respname <- as.character(formula[2])
  form <- paste0(respname, " ~ ", covform)
  f2 <- eval(parse(text=form))
  environment(f2) <- environment(formula)
  ## names of variables in the data, not the formula, with functions such as factor() stripped
  ## used for error message with incomplete "newdata" in summary()
```

Code

```
  covnames.bare <- unlist(lapply(forms, function(x)all.vars(delete.response(terms(x)))))
  attr(f2, "covnames") <- covnames.bare
  attr(f2, "covnames.orig") <- covnames
  f2
}

## User-supplied initial value functions don't have to include all
## four possible arguments: this expands them if they don't

expand.inits.args <- function(inits){
  inits2 <- inits
  formals(inits2) <- alist(t=,mf=,mml=,aux=)
  body(inits2) <- body(inits)
  inits2
}

## User-supplied summary output functions don't have to include all
## two possible arguments: this expands them if they don't

expand.summfn.args <- function(summfn){
  summfn2 <- summfn
  args <- c(alist(t=,start=), formals(summfn))
  formals(summfn2) <- args[!duplicated(names(args))]
  body(summfn2) <- body(summfn)
  summfn2
}

### On entry:
### event (status=1)           time1=event time
### right-censoring (status=0)  time1=lower bound
### left-censoring (status=2)   time1=upper bound
### interval-censoring (status=3)  time1=lower, time2=upper

### On exit
### time1=lower bound or event time
### time2=upper bound
### start=left truncation time
### so meaning of time1,time2 reversed with left-censoring

check.flexsurv.response <- function(Y){
  if (!inherits(Y, "Surv"))
    stop("Response must be a survival object")
  ### convert Y from Surv object to numeric matrix
  ### though "time" only used for initial values, printed time at risk, empirical hazard
  if (attr(Y, "type") == "counting")
    Y <- cbind(Y, time=Y[,"stop"] - Y[,"start"], time1=Y[,"stop"], time2=Inf)
  else if (attr(Y, "type") == "interval"){
    Y[,"time2"][Y[,"status"]==0] <- Inf   # upper bound with right censoring
    Y[,"time2"][Y[,"status"]==2] <- Y[,"time1"][Y[,"status"]==2]
    Y[,"time1"][Y[,"status"]==2] <- 0  #
    Y <- cbind(Y, start=0, stop=Y[,"time1"], time=Y[,"time1"])
  }
  else if (attr(Y, "type") == "right")
    Y <- cbind(Y, start=0, stop=Y[,"time"], time1=Y[,"time"], time2=Inf)
  else stop("Survival object type \"", attr(Y, "type"), "\"", " not supported")
  if (any(Y[,"time1"]<0)){
    stop("Negative survival times in the data")
  }
```

Code

```
  Y
}

compress.model.matrices <- function(mml){
  cbind.drop.intercept <- function(...)do.call("cbind", lapply(list(...), function(x)x[,-1,drop=FALSE]))
  X <- do.call("cbind.drop.intercept",mml)
  loc.cnames <- colnames(mml[[1]])[-1]
  anc.cnames <- unlist(mapply(function(x,y)sprintf("%s(%s)",x,y), names(mml[-1]), lapply(mml[-1],
function(x)colnames(x)[-1])))
  cnames <- c(loc.cnames, anc.cnames)
  colnames(X) <- cnames
  X
}

form.dp <- function(dlist, dfns, integ.opts){

  ## TODO check for format of dfn (args x, log)
  ## FIXME bug if object called d is found in global env
  ## check for existence in current frame.  inherits false?

  name <- dlist$name
  hname <- paste0("h",name); Hname <- paste0("H",name)
  dname <- paste0("d",name); pname <- paste0("p",name)
  rmstname <- paste0("rmst_",name)
  meanname <- paste0("mean_",name)
  qname <- paste0("q",name)
  rname <- paste0("r",name)
  if (is.function(dfns$d)) d <- dfns$d
  if (is.function(dfns$p)) p <- dfns$p
  if (is.function(dfns$h)) h <- dfns$h
  if (is.function(dfns$H)) H <- dfns$H
  if (is.function(dfns$r)) r <- dfns$r
  if (is.function(dfns$q)) q <- dfns$q
  if (is.function(dfns$mean)) meanf <- dfns$mean
  if (is.function(dfns$rmst)) rmst <- dfns$rmst
  if (!exists("h", inherits=FALSE)){
    if (exists(hname)) h <- get(hname)
    else {
      if (!exists("d")){
        if (exists(dname)) d <- get(dname)
        else stop("Neither density function \"",dname,
                  "\" nor hazard function \"", hname, "\" found")
      }
      if (!exists("p")){
        if (exists(pname)) p <- get(pname)
        else {
          message("Forming cumulative distribution function...")
          p <- integrate.dh(d, dlist, integ.opts, what="density")
        }
      }
      h <- function(x, ...){
        d(x,...)/(1 - p(x,...))
      }
    }
  }
  if (!exists("H", inherits=FALSE)){
    if (exists(Hname)) H <- get(Hname)
    else {
```

Code

```
    if (!exists("p")) { if (exists(pname)) p <- get(pname) }
    if (exists("p")){
      H <- function(x, ...){
        -log(1 - p(x, ...))
      }
    } else {
      message("Forming integrated hazard function...")
      H <- integrate.dh(h, dlist, integ.opts, what="hazard")
    }
  }
}
if (!exists("p", inherits=FALSE)){
  if (exists(pname)) p <- get(pname)
  else {
    p <- function(q, ...) {
      ret <- 1 - exp(-H(q, ...))
      #             ret[q==Inf] <- 1 # should have been handled already in cum.fn
      #             ret[q==0] <- 0
      ret
      ### TODO special values in other functions
    }
  }
}
if (!exists("q", inherits=FALSE)){
  if (exists(qname)) q <- get(qname)
  else {
    # giving this another name to avoid scoping issues
    # w/ name p also being an argument to q functions
    pfun <- p
    q <- function(p, ...) qgeneric(pfun, p)
  }
}
if (!exists("d", inherits=FALSE)){
  if (exists(dname)) d <- get(dname)
  else {
    d <- function(x, log=FALSE, ...) {
      if (log)
        log(h(x,...)) + log(1 - p(x, ...))
      else h(x,...) * (1 - p(x, ...))
    }
  }
}
if (!exists("rmst", inherits=FALSE)){
  if (exists(rmstname)) rmst <- get(rmstname)
  else {
    message("Forming integrated rmst function...")
    rmst <- function(t, start=0, ...) rmst_generic(p, t=t, start=start, ...)
  }
}
if (!exists("meanf", inherits=FALSE)){
  if (exists(rmstname)) meanf <- get(meanname)
  else {
    message("Forming integrated mean function...")
    meanf <- function(start=0, ...) rmst(t=Inf, start=start, ...)
  }
}
if (!exists("r", inherits=FALSE)){
  if (exists(rname)) r <- get(rname)
```

Code

```
    else r <- NULL
    ## random sampling function is currently only used for multi-state models
  }
  ## Check for existence of derivative functions
  ## conventionally called DLd, DLs
  ## if dfns$deriv set to FALSE on entry, derivatives not available
  if (is.function(dfns$DLd)) DLd <- dfns$DLd
  else if (is.null(dfns$deriv) && exists(paste0("DLd",name)))
    DLd <- get(paste0("DLd",name))
  else DLd <- NULL
  if (is.function(dfns$DLS)) DLS <- dfns$DLS
  else if (is.null(dfns$deriv) && exists(paste0("DLS",name)))
    DLS <- get(paste0("DLS",name))
  else DLS <- NULL

  list(p=p, d=d, h=h, H=H, r=r, DLd=DLd, DLS=DLS, rmst=rmst, mean= meanf,
       q=q, deriv = !is.null(DLd) && !is.null(DLS))
}


## Produce cumulative version of hazard function or density function
## by numerical integration

integrate.dh <- function(fn, dlist, integ.opts, what="dens"){

  cum.fn <- function(q, ...){
    args <- list(...)
    pars <- as.list(dlist$pars)
    names(pars) <- dlist$pars
    args.done <- numeric()
    ## if argument is unnamed, assume it is supplied in the default order
    for (i in seq(along=dlist$pars)){
      if(any(names(args)==dlist$pars[i])) {
        pars[[i]] <- args[[dlist$pars[i]]]
        args.done <- c(args.done, match(dlist$pars[i], names(args)))
      } else {
        pars[[i]] <- args[[i]]
        args.done <- c(args.done, i)
      }
    }
    ## any auxiliary arguments not in main distribution parameters
    rest <- args[setdiff(seq_along(args), args.done)]
    ## replicate all arguments to have the length of the longest one (=n)
    n <- max(sapply(c(list(q),pars), length))
    q <- rep(q, length=n)
    for (i in seq_along(pars)) pars[[i]] <- rep(pars[[i]], length=n)
    ret <- numeric(n)
    du <- function(u, ...)fn(u,...)
    ## then return a vector of length n
    for (i in 1:n){
      parsi <- lapply(pars, function(x)x[i])
      int.args <- c(list(f=du, lower=0, upper=q[i]), parsi, rest, integ.opts)
      if (q[i]==0) ret[i] <- 0
      else if (q[i]==Inf) {
        if (what=="density") ret[i] <- 1
        else if (what=="hazard") ret[i] <- Inf
      }
      else {
```

Code

```
      int <- try(do.call("integrate", int.args))
      #               if (inherits(int, "try-error")) browser()
      ret[i] <- int$value
    }
   }
   ret
  }

  cum.fn
}

flexsurvreg.edit <- function (formula, anc = NULL, data, weights, bhazard, subset,
                          na.action, dist, inits, fixedpars = NULL, dfns = NULL, aux = NULL,
                          cl = 0.95, integ.opts = NULL, sr.control = survreg.control(),
                          ...)
{
  call <- match.call()
  if (missing(dist))
    stop("Distribution \"dist\" not specified")
  if (is.character(dist)) {
    dist <- match.arg(tolower(dist), tolower(names(flexsurv.dists)))
    dist <- names(flexsurv.dists)[match(dist, tolower(names(flexsurv.dists)))]
    dlist <- flexsurv.dists[[dist]]
  }
  else if (is.list(dist)) {
    dlist <- check.dlist(dist)
  }
  else stop("\"dist\" should be a string for a built-in distribution, or a list for a custom distribution")
  dfns <- form.dp(dlist, dfns, integ.opts)
  parnames <- dlist$pars
  ancnames <- setdiff(parnames, dlist$location)
  check.formula(formula, dlist)
  if (is.null(anc)) {
    anc <- vector(mode = "list", length = length(ancnames))
    names(anc) <- ancnames
    for (i in ancnames) {
      anc[[i]] <- ancpar.formula(formula, i)
    }
  }
  else {
    if (!is.list(anc) || !all(sapply(anc, function(x) inherits(x,
                                                    "formula"))))
      stop("\"anc\" must be a list of formulae")
  }
  forms <- c(location = get.locform(formula, ancnames), anc)
  names(forms)[[1]] <- dlist$location
  indx <- match(c("formula", "data", "weights", "bhazard",
                  "subset", "na.action"), names(call), nomatch = 0)
  if (indx[1] == 0)
    stop("A \"formula\" argument is required")
  temp <- call[c(1, indx)]
  temp[[1]] <- as.name("model.frame")
  f2 <- concat.formulae(formula, forms)
  temp[["formula"]] <- f2
  if (missing(data))
    temp[["data"]] <- environment(formula)
  m <- eval(temp, parent.frame())
  m <- droplevels(m)
```

Code

```
  attr(m, "covnames") <- attr(f2, "covnames")
  attr(m, "covnames.orig") <- intersect(colnames(m), attr(f2,
                                                "covnames.orig"))
  Y <- check.flexsurv.response(model.extract(m, "response"))
  mml <- mx <- vector(mode = "list", length = length(dlist$pars))
  names(mml) <- names(mx) <- c(dlist$location, setdiff(dlist$pars,
                                               dlist$location))
  for (i in names(forms)) {
    mml[[i]] <- model.matrix(forms[[i]], m)
    mx[[i]] <- length(unlist(mx)) + seq_len(ncol(mml[[i]][,
                                                -1, drop = FALSE]))
  }
  X <- compress.model.matrices(mml)
  weights <- model.extract(m, "weights")
  if (is.null(weights))
    weights <- m$"(weights)" <- rep(1, nrow(X))
  bhazard <- model.extract(m, "bhazard")
  if (is.null(bhazard))
    bhazard <- rep(0, nrow(X))
  dat <- list(Y = Y, m = m, mml = mml)
  ncovs <- length(attr(m, "covnames.orig"))
  ncoveffs <- ncol(X)
  nbpars <- length(parnames)
  npars <- nbpars + ncoveffs
  if (missing(inits) && is.null(dlist$inits))
    stop("\"inits\" not supplied, and no function to estimate them found in the custom distribution list")
  if (missing(inits) || any(is.na(inits))) {
    yy <- ifelse(Y[, "status"] == 3 & is.finite(Y[, "time2"]),
                 (Y[, "time1"] + Y[, "time2"])/2, Y[, "time"])
    wt <- yy * weights * length(yy)/sum(weights)
    dlist$inits <- expand.inits.args(dlist$inits)
    inits.aux <- c(aux, list(forms = forms, data = if (missing(data)) NULL else data,
                             weights = temp$weights, control = sr.control, counting = (attr(model.extract(m,
                                                                                                           "response"), "type") ==
"counting")))
    auto.inits <- dlist$inits(t = wt, mf = m, mml = mml,
                              aux = inits.aux)
    if (!missing(inits) && any(is.na(inits)))
      inits[is.na(inits)] <- auto.inits[is.na(inits)]
    else inits <- auto.inits
  }
  if (!is.numeric(inits))
    stop("initial values must be a numeric vector")
  nin <- length(inits)
  if (nin < npars && ncoveffs > 0)
    inits <- c(inits, rep(0, length = npars - nin))
  else if (nin > npars) {
    inits <- inits[1:npars]
    warning("Initial values are a vector length > ", npars,
            ": using only the first ", npars)
  }
  else if (nin < nbpars) {
    stop("Initial values are a vector length ", nin, ", but distribution has ",
         nbpars, " parameters")
  }
  for (i in 1:nbpars) inits[i] <- dlist$transforms[[i]](inits[i])
  outofrange <- which(is.nan(inits) | is.infinite(inits))
  if (any(outofrange)) {
```

Code

```
    plural <- if (length(outofrange) > 1)
      "s"
    else ""
    stop("Initial value", plural, " for parameter", plural,
         " ", paste(outofrange, collapse = ","), " out of range")
  }
  names(inits) <- c(parnames, colnames(X))
  if (!is.null(fixedpars) && !is.logical(fixedpars) && (!is.numeric(fixedpars) ||
                                          any(!(fixedpars %in% 1:npars)))) {
    dots <- if (npars > 2)
      "...,"
    else ""
    stop("fixedpars must be TRUE/FALSE or a vector of indices in 1,",
         dots, npars)
  }
  if ((is.logical(fixedpars) && fixedpars == TRUE) || (is.numeric(fixedpars) &&
                                         identical(fixedpars, 1:npars))) {
    minusloglik <- minusloglik.flexsurv(inits, Y = Y, X = X,
                                weights = weights, bhazard = bhazard, dlist = dlist,
                                inits = inits, dfns = dfns, aux = aux, mx = mx)
    res.t <- matrix(inits, ncol = 1)
    inits.nat <- inits
    for (i in 1:nbpars) inits.nat[i] <- dlist$inv.transforms[[i]](inits[i])
    res <- matrix(inits.nat, ncol = 1)
    print(res)
    dimnames(res) <- dimnames(res.t) <- list(names(inits),
                                     "est")
    ret <- list(res = res, res.t = res.t, npars = 0, loglik = -as.vector(minusloglik),
                logliki = attr(minusloglik, "indiv"))
  }
  else {
    optpars <- inits[setdiff(1:npars, fixedpars)]
    optim.args <- list(...)
    if (is.null(optim.args$method)) {
      optim.args$method <- "BFGS"
    }
    gr <- if (dfns$deriv)
      Dminusloglik.flexsurv
    else NULL
    optim.args <- c(optim.args, list(par = optpars, fn = logLikFactory(Y = Y,
                                                    X = X, weights = weights, bhazard = bhazard, inits =
inits,
                                                    dlist = dlist, dfns = dfns, aux = aux, mx = mx,
                                                    fixedpars = fixedpars), gr = gr, Y = Y, X = X, weights =
weights,
                                bhazard = bhazard, dlist = dlist, inits = inits,
                                dfns = dfns, aux = aux, mx = mx, fixedpars = fixedpars,
                                hessian = TRUE))
    opt <- do.call("optim", optim.args)
    est <- opt$par
    if (all(!is.na(opt$hessian)) && all(!is.nan(opt$hessian)) &&
        all(is.finite(opt$hessian)) && all(eigen(opt$hessian)$values >
                                       0)) {
      opt$hessian <- opt$hessian + replicate(length(opt$par),
                                     abs(rnorm(n = length(opt$par),
                                               mean = 0.00001, sd = 0.00001)))
      cov <- solve(opt$hessian)
      se <- sqrt(diag(cov))
```

Code

```
    if (!is.numeric(cl) || length(cl) > 1 || !(cl >
                                        0) || !(cl < 1))
      stop("cl must be a number in (0,1)")
    lcl <- est - qnorm(1 - (1 - cl)/2) * se
    ucl <- est + qnorm(1 - (1 - cl)/2) * se
  }
  else {
    opt$hessian <- opt$hessian + replicate(length(opt$par),
                                    abs(rnorm(n = length(opt$par),
                                              mean = 0.00001, sd = 0.00001)))
    cov <- solve(opt$hessian)
    se <- sqrt(diag(cov))
    if (!is.numeric(cl) || length(cl) > 1 || !(cl >
                                        0) || !(cl < 1))
      stop("cl must be a number in (0,1)")
    lcl <- est - qnorm(1 - (1 - cl)/2) * se
    ucl <- est + qnorm(1 - (1 - cl)/2) * se
  }
  res <- cbind(est = inits, lcl = NA, ucl = NA, se = NA)
  res[setdiff(1:npars, fixedpars), ] <- cbind(est, lcl,
                                      ucl, se)
  colnames(res) <- c("est", paste(c("L", "U"), round(cl *
                                          100), "%", sep = ""), "se")
  res.t <- res
  for (i in 1:nbpars) {
    res[i, 1:3] <- dlist$inv.transforms[[i]](res[i,
                                        1:3])
    if (identical(body(dlist$transforms[[i]]), body(log)))
      res[i, "se"] <- exp(res.t[i, "est"]) * res.t[i,
                                          "se"]
    else if (identical(body(dlist$transforms[[i]]),
                       body(logh)))
      res[i, "se"] <- dexph(res.t[i, "est"]) * res.t[i,
                                          "se"]
    else if (!identical(dlist$transforms[[i]], identity))
      res[i, "se"] <- NA
  }
  minusloglik <- minusloglik.flexsurv(res.t[, "est"],
                                Y = Y, X = X, weights = weights, bhazard = bhazard,
                                dlist = dlist, inits = inits, dfns = dfns, aux = aux,
                                mx = mx)
  ret <- list(res = res, res.t = res.t, cov = cov, coefficients = res.t[,
                                                          "est"], npars = length(est), fixedpars = fixedpars,
              optpars = setdiff(1:npars, fixedpars), loglik = -opt$value,
              logliki = attr(minusloglik, "indiv"), cl = cl, opt = opt)
}
ret <- c(list(call = call, dlist = dlist, aux = aux, ncovs = ncovs,
              ncoveffs = ncoveffs, mx = mx, basepars = 1:nbpars, covpars = if (ncoveffs >
                                                          0) (nbpars + 1):npars else NULL, AIC = -2 *
ret$loglik +
                2 * ret$npars, data = dat, datameans = colMeans(X),
              N = nrow(dat$Y), events = sum(dat$Y[, "status"] == 1),
              trisk = sum(dat$Y[, "time"]), concat.formula = f2, all.formulae = forms,
              dfns = dfns), ret)
if (isTRUE(getOption("flexsurv.test.analytic.derivatives")) &&
    (dfns$deriv)) {
  if (is.logical(fixedpars) && fixedpars == TRUE) {
    optpars <- inits
```

Code

```
      fixedpars = FALSE
    }
    ret$deriv.test <- deriv.test(optpars, Y, X, weights,
                                 bhazard, dlist, inits, dfns, aux, mx, fixedpars)
  }
  class(ret) <- "flexsurvreg"
  ret
}

svysurvreg.survey.design<-
  function (formula, design, dist, weights=NULL, subset=NULL, ...)
  {
    subset <- substitute(subset)
    subset <- eval(subset, model.frame(design), parent.frame())
    data <- model.frame(design)
    g <- match.call()
    g$formula <- eval.parent(g$formula)
    g$design <- NULL
    g$var <- NULL
    if (is.null(g$weights))
      g$weights <- quote(.survey.prob.weights)
    else g$weights <- bquote(.survey.prob.weights * .(g$weights))
    g[[1]] <- quote(survreg)
    g$formula <- formula
    g$data <- quote(data)
    g$subset <- quote(.survey.prob.weights > 0)
    g$model <- TRUE
    data$.survey.prob.weights <- (1/design$prob)/mean(1/design$prob)
    if (!all(all.vars(formula) %in% names(data)))
      stop("all variables must be in design= argument")
    g <- with(list(data = data), eval(g))
    g$call <- match.call()
    #g$call[[1]] <- as.name(.Generic)
    # g$printcall <- sys.call(-1)
    #g$printcall[[1]] <- as.name(.Generic)
    class(g) <- c("svysurvreg", class(g))
    g$survey.design <- design
    nas <- g$na.action
    if (length(nas))
      design <- design[-nas, ]
    dbeta.subset <- resid(g, "dfbeta", weighted = TRUE)
    if (nrow(design) == NROW(dbeta.subset)) {
      dbeta <- as.matrix(dbeta.subset)
    }
    else {
      dbeta <- matrix(0, ncol = NCOL(dbeta.subset), nrow = nrow(design))
      dbeta[is.finite(design$prob), ] <- dbeta.subset
    }
    g$inv.info <- g$var
    if (inherits(design, "survey.design2"))
      g$var <- svyrecvar(dbeta, design$cluster, design$strata,
                         design$fpc, postStrata = design$postStrata)
    else if (inherits(design, "twophase"))
      g$var <- twophasevar(dbeta, design)
    else if (inherits(design, "twophase2"))
      g$var <- twophase2var(dbeta, design)
    else if (inherits(design, "pps"))
      g$var <- ppsvar(dbeta, design)
```

Code

```
    else g$var <- svyCprod(dbeta, design$strata, design$cluster[[1]],
                           design$fpc, design$nPSU, design$certainty, design$postStrata)
    g$ll <- g$loglik
    g$loglik <- g$ll
    g$degf.resid <- degf(design) - length(coef(g)[!is.na(coef(g))]) +
      1
    g
  }

residuals.flexsurv <- function(object, type=c('response', 'deviance',
                                       'dfbeta', 'dfbetas', 'working', 'ldcase',
                                       'ldresp', 'ldshape', 'matrix'),
                         rsigma =TRUE, collapse=FALSE, weighted=FALSE, ...) {
  type <-match.arg(type)
  n <- object$N
  weights <- object$weights
  # vv <- object$var
  if (!is.null(dd$dist)) {
    dd <- survreg.distributions[[dd$dist]]
  }
  deviance <- dd$deviance
  dens <- dd$density
  status <- y[,ncol(y)]
  eta <- object$linear.predictors
  z <- (y[,1] - eta)/sigma
  dmat <- dens(z, object$parms)
  dtemp<- dmat[,3] * dmat[,4]    #f'
  if (any(status==3)) {
    z2 <- (y[,2] - eta)/sigma
    dmat2 <- dens(z2, object$parms)
  }
  else {
    dmat2 <- dmat   #dummy values
    z2 <- 0
  }
  tdenom <- ((status==0) * dmat[,2]) +  #right censored
    ((status==1) * 1 )       +  #exact
    ((status==2) * dmat[,1]) +  #left
    ((status==3) * ifelse(z>0, dmat[,2]-dmat2[,2],
                          dmat2[,1] - dmat[,1])) #interval
  g <- log(ifelse(status==1, dmat[,3]/sigma, tdenom))  #loglik
  tdenom <- 1/tdenom
  dg <- -(tdenom/sigma) *(((status==0) * (0-dmat[,3])) +    #dg/ eta
                          ((status==1) * dmat[,4]) +
                          ((status==2) * dmat[,3]) +
                          ((status==3) * (dmat2[,3]- dmat[,3])))

  ddg <- (tdenom/sigma^2) *(((status==0) * (0- dtemp)) +  #ddg/eta^2
                            ((status==1) * dmat[,5]) +
                            ((status==2) * dtemp) +
                            ((status==3) * (dmat2[,3]*dmat2[,4] - dtemp)))
  ds  <- ifelse(status<3, dg * sigma * z,
                tdenom*(z2*dmat2[,3] - z*dmat[,3]))
  dds <- ifelse(status<3, ddg* (sigma*z)^2,
                tdenom*(z2*z2*dmat2[,3]*dmat2[,4] -
                          z * z*dmat[,3] * dmat[,4]))
  dsg <- ifelse(status<3, ddg* sigma*z,
                tdenom *(z2*dmat2[,3]*dmat2[,4] - z*dtemp))
```

Code

```
  deriv <- cbind(g, dg, ddg=ddg- dg^2,
                 ds = ifelse(status==1, ds-1, ds),
                 dds=dds - ds*(1+ds),
                 dsg=dsg - dg*(1+ds))
  if (type=='dfbeta') {
    score <- deriv[,2] * x  # score residuals
    if (rsigma) {
      if (nstrata > 1) {
        d4 <- matrix(0., nrow=n, ncol=nstrata)
        d4[cbind(1:n, strata)] <- deriv[,4]
        score <- cbind(score, d4)
      }
      else score <- cbind(score, deriv[,4])
    }
    rr <- score %*% vv
  }
  if (weighted==TRUE) {
    rr <- rr * weights
  }
  rr
}

svyflexsurvreg.survey.design<-
  function (formula, design, dist, weights=NULL, subset=NULL, ...)
  {
    subset <- substitute(subset)
    subset <- eval(subset, model.frame(design), parent.frame())
    data <- model.frame(design)
    g <- match.call()
    g$formula <- eval.parent(g$formula)
    g$design <- NULL
    g$var <- NULL
    if (is.null(g$weights))
      g$weights <- quote(.survey.prob.weights)
    else g$weights <- bquote(.survey.prob.weights * .(g$weights))
    g[[1]] <- quote(flexsurvreg.edit)
    g$formula <- formula
    g$data <- quote(data)
    g$subset <- quote(.survey.prob.weights > 0)
    g$dist <- dist
    data$.survey.prob.weights <- (1/design$prob)/mean(1/design$prob)
    if (!all(all.vars(formula) %in% names(data)))
      stop("all variables must be in design= argument")
    g <- with(list(data = data), eval(g))
    g$call <- match.call()
    #g$call[[1]] <- as.name(.Generic)
    # g$printcall <- sys.call(-1)
    #g$printcall[[1]] <- as.name(.Generic)
    class(g) <- c("svysurvreg", class(g))
    g$survey.design <- design
    nas <- g$na.action
    if (length(nas))
      design <- design[-nas, ]
    # dbeta.subset <- resid(g, "dfbeta", weighted = TRUE)
    # if (nrow(design) == NROW(dbeta.subset)) {
    #   dbeta <- as.matrix(dbeta.subset)
    # }
    # else {
```

Code

```
  #   dbeta <- matrix(0, ncol = NCOL(dbeta.subset), nrow = nrow(design))
  #   dbeta[is.finite(design$prob), ] <- dbeta.subset
  # }
  # g$inv.info <- g$var
  # if (inherits(design, "survey.design2"))
  #   g$var <- svyrecvar(dbeta, design$cluster, design$strata,
  #                      design$fpc, postStrata = design$postStrata)
  # else if (inherits(design, "twophase"))
  #   g$var <- twophasevar(dbeta, design)
  # else if (inherits(design, "twophase2"))
  #   g$var <- twophase2var(dbeta, design)
  # else if (inherits(design, "pps"))
  #   g$var <- ppsvar(dbeta, design)
  # else g$var <- svyCprod(dbeta, design$strata, design$cluster[[1]],
  #                        design$fpc, design$nPSU, design$certainty, design$postStrata)
  g$ll <- g$loglik
  g$loglik <- g$ll
  g$degf.resid <- degf(design) - length(coef(g)[!is.na(coef(g))]) +
    1
  g
}
```

The coefficients for the models used in CS, Section B.3.10.3, page 77, i.e., the unanchored MAIC, are shown in Table . Table also contains coefficients used in ITC scenarios reported in clarification question B6.

Table 12: ITC parametric model coefficients

Scenario	Model	Coefficient	Survival function MS EXCEL parameterisation
Anchored MAIC constant HR (Weibull HR base case)	Cabozantinib OS	mu = 2.57 sigma = 0.85 gamma = 0.27	Weibull S(t)=EXP(-EXP(-mu/sigma-gamma/sigma)*t^(1/sigma))
	Regorafenib OS	Regorafenib vs. RESORCE placebo HR = 0.67 Therefore mu = 2.91 sigma = 0.85	Weibull S(t)=EXP(-EXP(-mu/sigma)*t^(1/sigma))
	Cabozantinib PFS	mu =1.25 sigma = 0.76 gamma = 0.79	Weibull S(t)=EXP(-EXP(-mu/sigma-gamma/sigma)*t^(1/sigma))
	Regorafenib PFS	Regorafenib vs. RESORCE placebo HR = 0.44 Therefore mu = 1.88 sigma = 0.76	Weibull S(t)=EXP(-EXP(-mu/sigma)*t^(1/sigma))
Anchored MAIC constant HR (Cox PH)	Cabozantinib OS	mu = 2.84 sigma = 0.85	Weibull S(t)=EXP(-EXP(-mu/sigma-gamma/sigma)*t^(1/sigma))
	Cabozantinib PFS	mu = 2.04 sigma = 0.78	Weibull S(t)=EXP(-EXP(-mu/sigma-gamma/sigma)*t^(1/sigma))
Anchored MAIC time-varying HR	Cabozantinib OS	mu = 2.43 sigma = 0.63	Log-logistic S(t)=1/(1+EXP(-(mu/sigma))*t^(1/sigma))
	Cabozantinib PFS	mu = 1.66 sigma = 0.52	Log-logistic S(t)=1/(1+EXP(-(mu/sigma))*t^(1/sigma))
Unanchored MAIC	Cabozantinib OS	mu = 2.43 sigma = 0.63	Log-logistic S(t)=1/(1+EXP(-(mu/sigma))*t^(1/sigma))
	Regorafenib OS	mu = 2.33 sigma = 0.61	Log-logistic S(t)=1/(1+EXP(-(mu/sigma))*t^(1/sigma))
	Cabozantinib PFS	mu = 1.87 sigma = 0.84 Q = 0.57	Generalised gamma S(t)=IF(Q>0,1,0)+IF(Q>0,-1,1)*IFERROR(GAMMA.DIST((Q^-2)*EXP(Q*((LN(t)-mu)/sigma)),Q^-2,1,TRUE),0)
	Regorafenib PFS	mu = 1.11 sigma = 0.93 Q = -0.34	Generalised gamma S(t)=IF(Q>0,1,0)+IF(Q>0,-1,1)*IFERROR(GAMMA.DIST((Q^-2)*EXP(Q*((LN(t)-mu)/sigma)),Q^-2,1,TRUE),0)

Abbreviations: HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; OS, overall survival; PFS, progression-free survival; PH, proportional hazard;

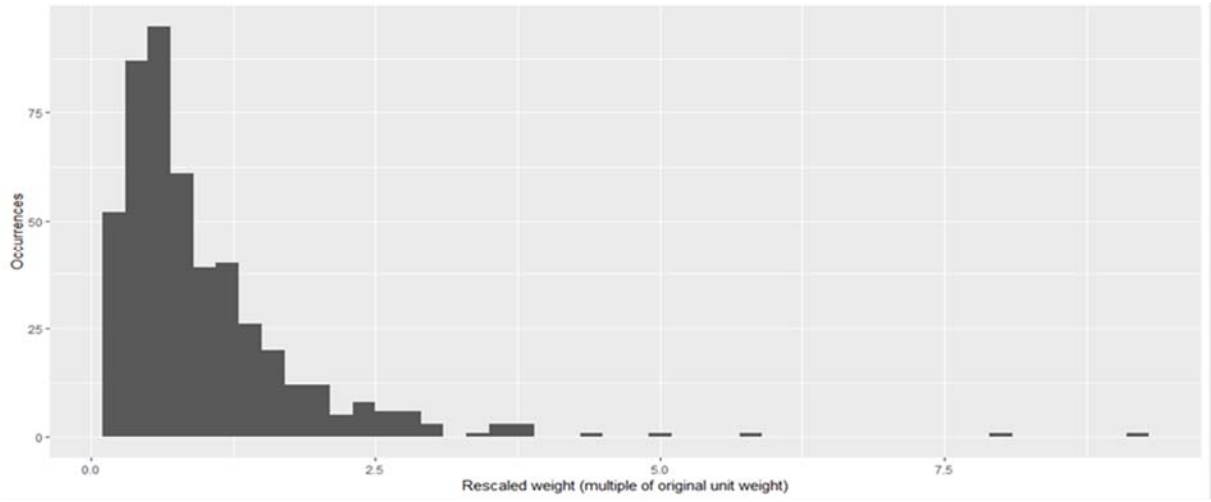
A23. CS, Section B.3.10.5, page 80. Please provide descriptive statistics about the rescaled weights obtained from the MAIC analysis (e.g. mean, median, Q1, Q2, Q3,

minimum and maximum). Please also provide a histogram of the distribution of the rescaled weights.

Company response

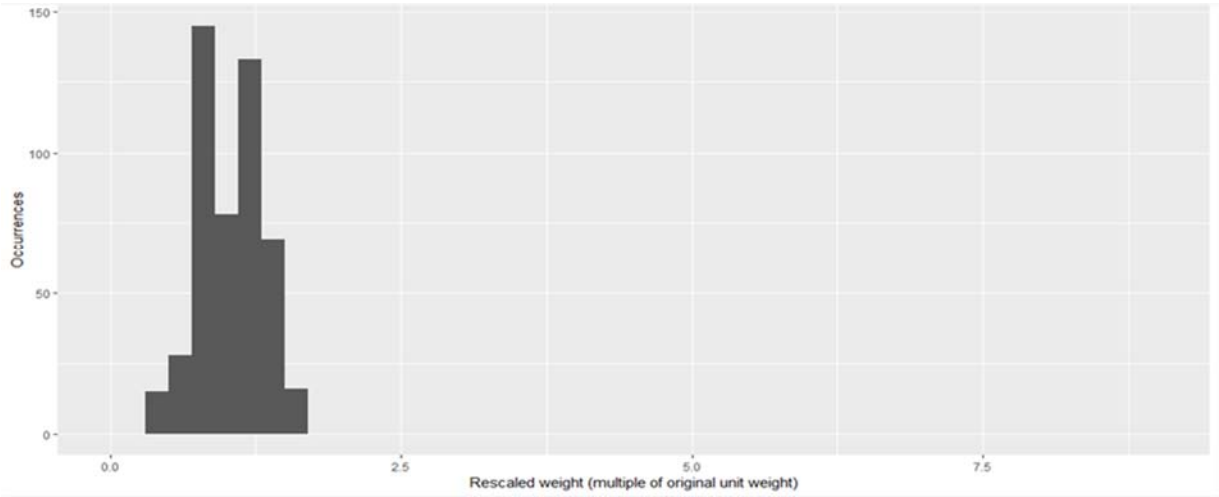
The distribution of the weights for Scenario 1 is examined in Figure 14, where the weights have been rescaled relative to the original unit weights of each individual.

Figure 14: Histogram of rescaled weights (Scenario 1)



The histogram in Figure 15 examines the distribution of rescaled weights for Scenario 2. The histogram for Scenario 1 (Figure 14) shows that there are some very large rescaled weights, with a maximum at 9.21. Scenario 2 reduces the presence of extreme weights (the maximum rescaled weight is 1.61), resulting in an approximate ESS which is very close to the original sample size and pulling the rescaled weights closer to one. Scenario 2 provides greater statistical power and precision than Scenario 1. However, Scenario 2 does not match some characteristics that are considered to be important effect modifiers by the clinical experts and which differ considerably across trials (e.g. duration of prior sorafenib treatment and geographical region). Also, the automatic variable selection method employed only evaluates the most contributory predictor variables for the primary survival endpoint, OS, and not for PFS or safety outcomes. In addition, the weighting of certain characteristics could drive the variables that have not been matched, moving the average for these variables further away from the values reported in RESORCE. However, this effect does not appear to be significant in the scenarios considered.

Figure 15: Histogram of rescaled weights (Scenario 2)



A24. CS, Section B.3.10.4, page 80. The CS states *"the ITC results suggest that cabozantinib has … similar tolerability compared to regorafenib".* However, the results of the Bucher ITC safety analysis (Table 26) give OR point estimates of 0.2 for hypertension and 0.6 for elevated aspartate aminotransferase. Please clarify the statement.

Company response

The p-values for hypertension and aspartate aminotransferase were not statistically significant. Only for diarrhoea is this clarification relevant as the p-value was statistically significant in the unanchored MAIC as regorafenib was associated with lower rates of diarrhoea in the RESORCE trial compared to cabozantinib in a pure second line population (weighted or unweighted). As mentioned in response to clarification A19 above, it is important to note that there may have been underreporting of certain adverse events and especially serious adverse events in HCC in the RESORCE trial since patients that discontinued prior sorafenib due to sorafenib-related toxicity were excluded from the study and sorafenib belongs to the same pharmacological class. The true impact on the reported safety profile for regorafenib in HCC is therefore unknown based on the published evidence that is available to-date. Further, clinical experts agree that the safety profile is similar between cabozantinib and regorafenib. Hence, the Company have stated in the submission that

the tolerability is the same between the two regimens, despite a statistically significant finding for an association between cabozantinib treatment and diarrhoea.

A25. CS, Appendix D, Section D.1.1.9, Table 14, page 30. This table summarises OS data for the RESORCE trial of regorafenib for data-cuts in 2016, 2017 and 2018 (HRs 0.63, 0.61 and 0.62, respectively). Please explain why the earlier data cut from 2016 has been used in the ITCs (CS, Section B.3.10.2, Table 24, page 62) rather than the most recent data cut.

Company response

The 2017 and 2018 OS data cuts do not report the OS KM curve for use in a population adjusted indirect comparison. The reported HRs show a small difference between data cuts; however, the Bucher ITC has been updated to incorporate the latest data cut and the results are reported in the response to clarification A13.

A26. Priority. CS, Appendix D, Section D.1.1.9, Table 14, page 30. This table summarises PFS data for the RESORCE trial using both mRECIST (HR 0.46) and RECIST 1.1 (HR 0.43). The CELESTIAL trial measures PFS using RECIST 1.1 (stated in CS, Section B.3.6.2, page 49). Therefore:

- Please explain why the Bucher ITC (CS, Section B.3.10.2, Table 24**,** page 62) uses PFS based on mRECIST rather than RECIST 1.1 for the RESORCE trial.
- In addition to the second-line Bucher ITC requested in Question A14, please also provide a second-line Bucher ITC for PFS based on RECIST 1.1 data for both CELESTIAL and RESORCE.

Company response

There is no explanation for the choice of mRECIST for RESORCE in the Bucher ITC; it seems to have been an oversight when extracting data from the trial publication. Regarding the results when using RECIST 1.1 for both RESORCE and CELESTIAL, please see response to clarification A13 (we assume the reference to A14 is a typo).

Section B: Clarification on cost comparison

B1. Priority. CS, Section B.4.2.1, page 85. The cost comparison assumes that both cabozantinib and regorafenib are given until progression and that PFS, which is used

as a proxy for time on treatment (ToT), is equivalent between the regimens. However, some patients in RESORCE and CELESTIAL continued to receive their allocated treatment beyond disease progression. Is there any evidence to support the assumption of equivalent ToT between the regimens?

Company response

In the regorafenib NICE appraisal (TA555) the clinical expert explained that 80% of patients would stop treatment on progression. The company agreed with the conclusion that most people would stop treatment if their disease progressed and provided a new survey which found that 8 of the 9 respondents would stop treatment at progression.

Figure 3 in clarification A11, shows a comparison of the cabozantinib PFS and TTD KM in the CELESTIAL trial. [redacted]. To assess the difference in time on treatment between using PFS and TTD, parametric models were fit to the cabozantinib TTD following the same process as for PFS. Figure 16 shows the parametric fits and Table shows the statistical fit. The generalised gamma and lognormal model had similar optimum AIC and BIC; however, the lognormal had a marginally better visual fit and so was selected as the base case. The TTD 15-year restricted mean was [redacted] months compared to the [redacted] months calculated using PFS.

Figure 16: TTD cabozantinib parametric fits

Abbreviations: KM, Kaplan-Meier; TTD, time to treatment discontinuation
Source: CELESTIAL individual patient level data

Table 13: AIC and BIC statistics for TTD parametric fits

Model	AIC	BIC
Exponential	1248.48	1252.63
Gompertz	1248.76	1257.05
Log-logistic	1194.34	1202.65
Lognormal	1179.61	**1187.90**
Weibull	1244.20	1252.49
Generalised gamma	**1177.38**	1189.81

Abbreviations: AIC, akaike information criterion; BIC, bayesian information criterion.; TTD, time to treatment discontinuation

B2. CS, Section B.3.3, page 33. Please provide a summary of subsequent anticancer treatments received in CELESTIAL. Please provide an equivalent summary for RESORCE, if this information is publicly available.

Company response

Table 14: Subsequent anticancer therapy in the CELESTIAL trial

Subsequent systemic anticancer therapy, n (%)	Cabozantinib (N = 470)	Placebo (N = 237)
Any non-radiation systemic or local liver-directed anticancer therapy, n (%)	123 (26)	78 (33)
Any systemic anticancer therapy	117 (25)	70 (30)
Sorafenib	19 (4)	4 (2)
Regorafenib	11 (2)	3 (1)
Anti-PD-1/PD-L1	8 (2)	15 (6)
Lenvatinib	1 (<1)	0
Cytotoxic chemotherapy	57 (12)	40 (17)
Investigational agent	28 (6)	17 (7)
Any non-radiation local liver-directed anticancer therapy	15 (3)	13 (5)

Abbreviations: *PD-1 programmed death-1; PD-L1, programmed death-ligand 1.*
Source: *Abou-Alfa et al., 2018.*

Twenty-five percent of patients continued treatment with a subsequent systemic anticancer therapy in the CELESTIAL trial. The most prevalent subsequent treatment was cytotoxic chemotherapy, followed by an anti-PD-1/PD-L1 inhibitor and either sorafenib or regorafenib.

To our knowledge the only publicly available information on subsequent anticancer treatments in the RESORCE trial is reported in a 2017 assessment report by the European Medicines Agency. Table 15 outlines the subsequent anticancer treatments for patients in the RESORCE trial.

Table 15: Systemic anti-cancer therapy in the RESORCE trial

ATC Classification Subclass WHO-DD Version 3q2005	Placebo N=194(100%)	Regorafenib 160mg N=379 (100%)	Total N=573(100%)
Number of subjects (%) with at least one medication	59 (30.4%)	88 (23.2%)	147 (25.7%)
Antineoplastic and Immunomodulating Agents	54 (27.8%)	76 (20.1%)	130 (22.7%)
Antineoplastic Agents	54 (27.8%)	73 (19.3%)	127 (22.2%)
Endocrine Therapy	1 (0.5%)	2 (0.5%)	3 (0.5%)
Immunostimulants	1 (0.5%)	2 (0.5%)	3 (0.5%)
Immunosuppressive Agents	2 (1.0%)	5 (1.3%)	7 (1.2%)

Source: EMA, 2017.

In the regorafenib arm 23.2% of the patients received subsequent systemic anticancer therapies. Details on the specific subsequent treatment that patients received is not given but is stated as antineoplastic and immunomodulating agents. The majority of

patients received an antineoplastic agent as a subsequent treatment. These may include cytotoxic chemotherapies as outlined in the subsequent treatments in CELESTIAL.

A key difference in subsequent treatments available during the RESORCE and CELESTIAL trials is the availability of subsequent treatments. The EMA assessment report for regorafenib was published in July 2017, 2 months before the first PD-1/PD-L1 inhibitor (nivolumab) was approved by the FDA for HCC in September 2017. Nivolumab was FDA approved at the time of reporting the subsequent treatments in the CELESTIAL trial. This may explain the difference in patients who subsequentially received anti PD-1/PD-L1 therapy in CELESTIAL compared to RESORCE.

B3. Priority. Given that relative dose intensity (RDI) was imperfect in both the regorafenib arm of RESORCE and the cabozantinib arm of CELESTIAL, please clarify why the base case analysis assumes 100% RDI.

Company response

In our cost comparison analysis, we presented two scenarios related to the relative dose intensity (RDI) of cabozantinib and regorafenib. These include our base case of 100% RDI (full pack dosing) and the RDI as reported in the CELESTIAL and RESORCE trials. The base case was informed by a previous regorafenib NICE submission (TA555).

It is noted that the applicant in TA555 presented an argument that showed drug wastage can be eliminated to reflect the RDI as reported in the trial. The NICE committee accepted that wastage could be reduced but not eliminated entirely. Overall, it was determined that the evidence to support the use of the trial reported RDI was significantly uncertain. As a result of this uncertainty the ERG presented two scenarios to calculate the annual cost of regorafenib: a pessimistic scenario where full pack dosing was assumed and an optimistic scenario where the trial reported RDI was used. The committee noted that the likely true cost in clinical practice is between the range of the two scenarios presented by the ERG.

Given the uncertainty shown regarding the use of the RDI in TA555, this submission presents the conservative scenario of assuming full pack dosing as the base case. In the sensitivity analysis a scenario is presented where the corresponding RDI's are

used. To account for the uncertainty in our assumption of drug wastage, a scenario is presented where RDI is 100% but no drug wastage is assumed. See section B4.4 for sensitivity and scenario analyses.

B4. Priority. CELESTIAL included the collection of EQ-5D data. Please provide EQ-5D utility values for second-line patients (a) for patients who are progression-free and (b) for patients with progressed disease. For each estimate, please provide the mean and 95% confidence interval. Please present these estimates by treatment group and for both treatment groups pooled.

Company response

A utility analysis was conducted using EQ-5D-3L data mapped from EQ-5D-5L data collected during the CELESTIAL study. The method published by van Hout, et al. was employed and UK value sets were used.

Three statistical models were employed to analyse the data:

- Ordinary Least Squares (OLS) regression - OLS model does not consider repeated EQ-5D-5L assessments for patients between study visits.
- Tobit regression with repeated measurements - The Tobit regression model has been previously used in other studies to derive utilities due to the presence of negative utility values (corresponding to health states worse than death). Using Tobit model negative utility values were transformed to 0.
- Mixed model for repeated measurements - Allows repeated EQ-5D-5L measurements at patient level to be considered given that patients provided several assessments during the study follow-up period.

The selection of the preferred model was defined based on the following criteria:

- Model reflecting the repeated nature of measurements
- Selection based on AIC measurements
- Smallest difference between the predicted and the observed values

The multivariable OLS regression model had the lowest AIC (-4391.09). However, this model does not reflect the nature of data collected given that it does not consider repeated measures of EQ-5D health states between study visits. Due to the repeated measures at patient level at different timepoints, this method was not considered optimal for the dataset. In addition, the number of questionnaires reported by each patient can be different and this can produce a bias in the results.

The multivariable Tobit regression with repeated measurements model was selected as the best option; it had a lower AIC compared to the mixed model for repeated measurements (-1772.93 vs -1931.16) and the errors obtained with prediction using the Tobit regression were closer to zero. This model considers that each patient has a different number of questionnaires but does require imputation in response variable, by imputing all negative utility values as zero. The summary of the results for the multivariable Tobit regression with mixed model for repeated measurements models are shown in Table 1 and Table .

Table 16: Summary of utility values for pooled treatment groups

Health state	Mean utility value	Standard error
Progression-free	[redacted]	[redacted]
Additional progressed disease disutility	[redacted]	[redacted]

Table 17: Summary of utility values for each treatment group

Health state	Mean utility value	Standard error
Progression-free	[redacted]	[redacted]
Additional progressed disease disutility	[redacted]	[redacted]
Additional cabozantinib arm disutility	[redacted]	[redacted]
Additional placebo arm disutility	[redacted]	-

B5. CC Model, worksheet “calculations”, cell D9. The model estimates that the net drug acquisition costs for regorafenib will be the same irrespective of whether wastage is included (net cost = [redacted]). Please confirm that this is due to the 1-week period off treatment at the end of each regorafenib treatment cycle.

Company response

Yes, this is correct. Based on the time on treatment of [redacted] months and a cycle length of 28 days for regorafenib, the average cycle length of regorafenib is [redacted] months. In the base case where wastage is assumed, to obtain the full [redacted] cycles of regorafenib 8 packs are required. No tablets in this scenario are wasted given that the last 7 days in the treatment cycle is off treatment.

In the scenario where no wastage is assumed the patient will still need 8 packs to fulfil [redacted] treatment cycles of regorafenib. Therefore, the annual cost of regorafenib in both scenarios is [redacted].

B6. Priority. Please fit standard parametric survival models to the OS data for the cabozantinib arm of the second-line subgroup in the CELESTIAL trial and apply the inverse HRs from the MAICs to estimate OS for the regorafenib arm. Please use NICE Technical Support Document 14 to guide the selection of the preferred OS model. Please present this analysis for the anchored MAICs (both time varying and constant HR). If time permits, please also extend this analysis to estimate net incremental QALYs for cabozantinib versus regorafenib using the data on PFS and EQ-5D collected in CELESTIAL.

Company response

The parametric fits for the population-adjusted cabozantinib OS and PFS data are shown in Figure 6 and Figure 10 respectively. The parametric fits for regorafenib OS and PFS from the RESORCE trial are shown in Figure 7 and Figure 11 respectively. This data is used in the unanchored MAIC scenario as described in CS, Section B.3.10.3, page 77. The following sections describe the other ITC scenarios and outcomes.

Anchored MAIC constant HR model fitting

The anchored MAIC using a constant HR was considered the base case as described in clarification A12. This scenario is conservative and does not require as strong assumptions as the unanchored MAIC, i.e., not all effect modifiers and prognostic variables need to be accounted for in an anchored comparison.

Anchored comparisons are compatible with a PH modelling approach however the HR needs to be applied to a base survival curve. Therefore, the following aspects must be

taken into consideration. Firstly, PH modelling is only compatible within PH models such as the exponential, Gompertz or Weibull. Log-logistic and log-normal models, for instance, are accelerated failure time models and do not produce a single HR. The PH assumption does not hold with these models. Secondly, the model type used to derive the HR must be the same as that fitted to the base survival curve. It is theoretically incorrect to apply a HR derived from a different parametric model or from a Cox PH model. Thus, the scenario was modelled using the following steps:

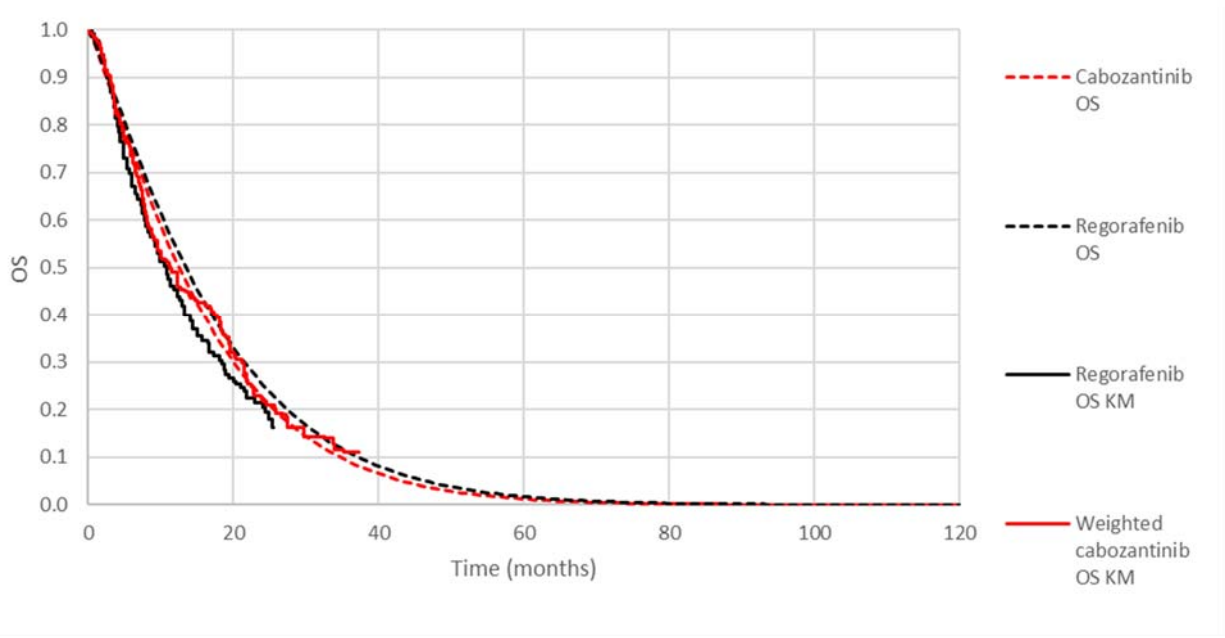
1. Fit a parametric model to the CELESTIAL data with treatment group as a covariate

2. Fit a parametric model to the RESORCE data with treatment group as a covariate

3. Apply the HR derived from Step 2 (the relative effect of regorafenib vs. placebo) to the placebo arm of CELESTIAL to derive a placebo-adjusted survival curve for regorafenib

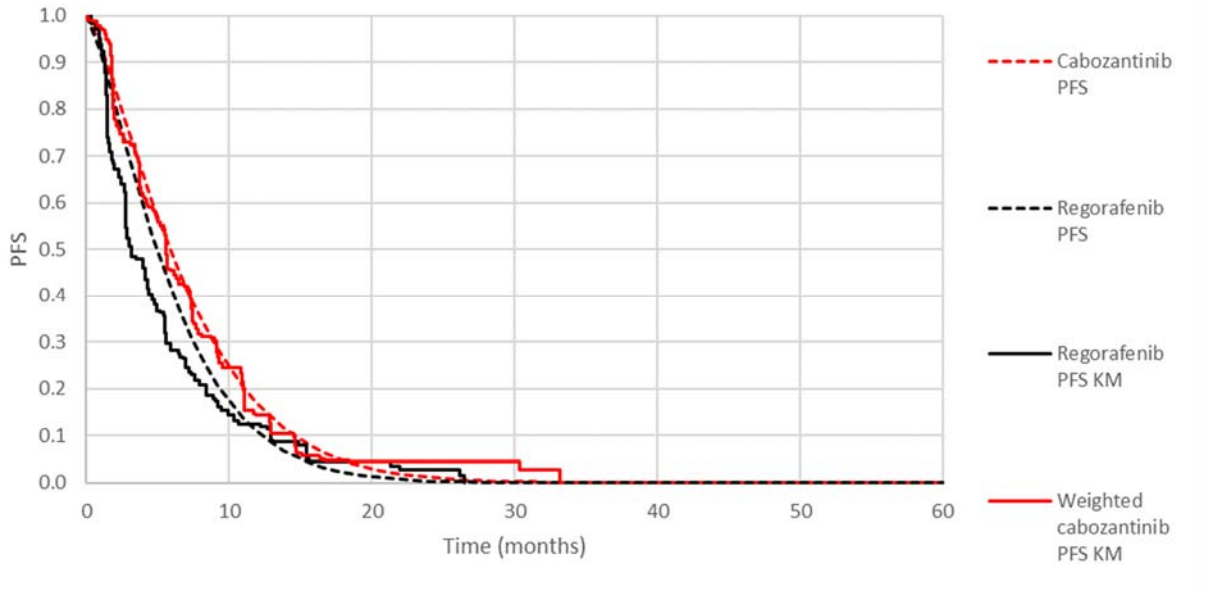
The Weibull distribution was selected to model survival, on the basis of lower AIC and BIC, indicating superior fit, and better fit to the observed Kaplan-Meier curves and observed log-cumulative hazards upon visual inspection. Cabozantinib and regorafenib OS are shown in Figure 17 and the PFS for both treatments are shown in Figure 18. The examination of the proportional hazards assumption in CS, Section B.3.10.3, page 72-75 showed that the use of a constant HR may not be appropriate for modelling both OS and PFS endpoints. This is illustrated by the modelled regorafenib OS and PFS, which generates greater estimates than the regorafenib KM observed in the RESORCE trial, biasing the comparison against cabozantinib.

Figure 17: Cabozantinib and regorafenib OS generated from anchored MAIC constant HR (Weibull HR)



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; OS, overall survival

Figure 18: Cabozantinib and regorafenib PFS generated from anchored MAIC constant HR (Weibull HR)

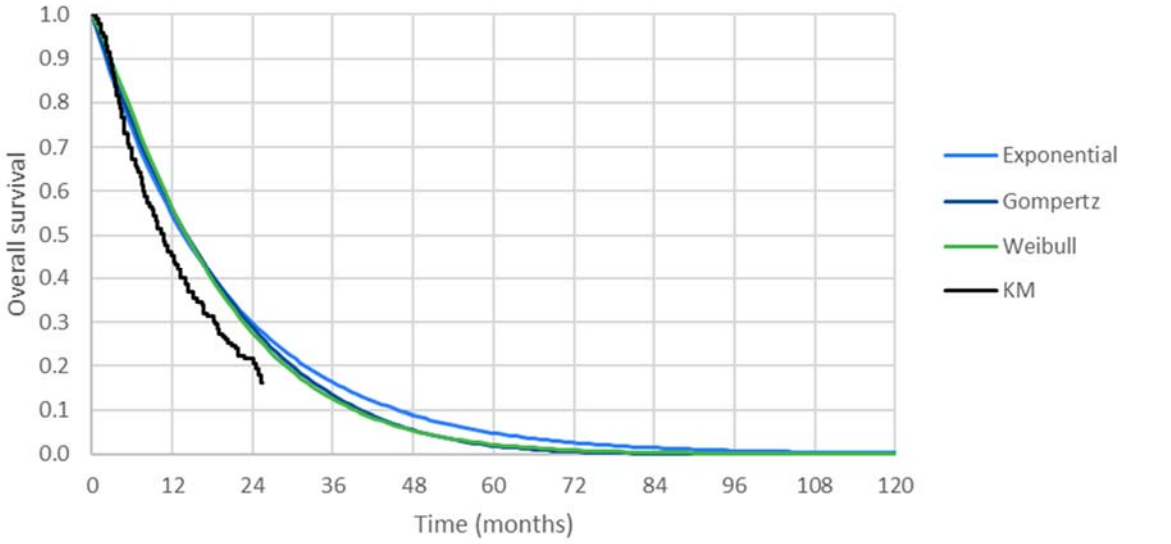


Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

For completion an anchored MAIC constant HR scenario using the Cox PH model as suggested by the ERG was also explored. Similarly to the Weibull HR scenario, model

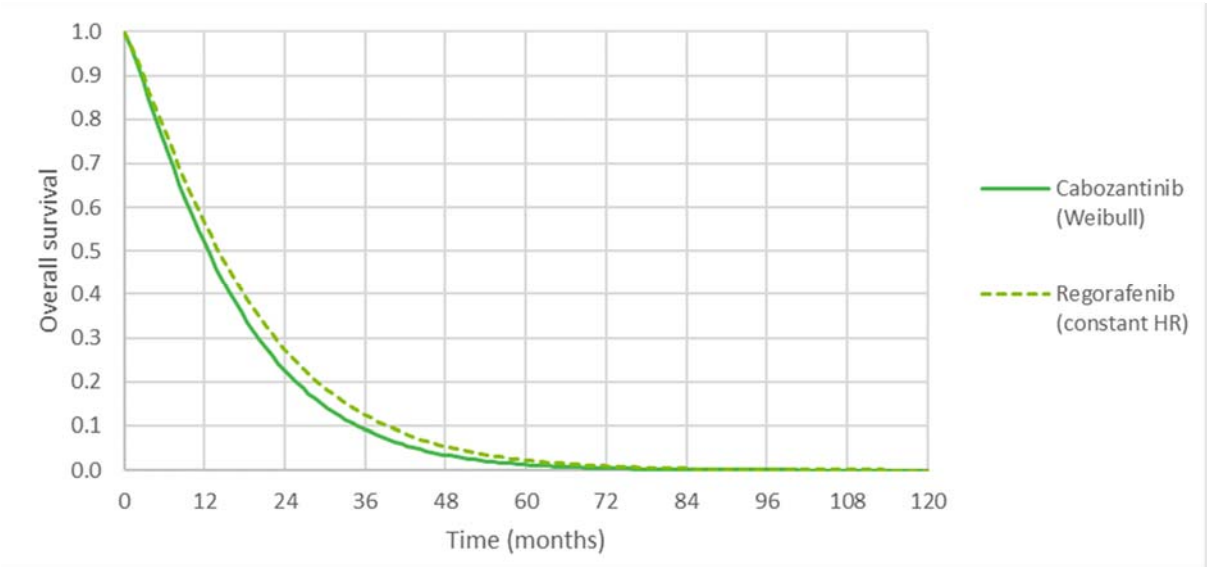
selection was restricted to only models that were compatible with proportional hazards and the Weibull was selected as the base case (Figure 20) for this scenario due to the statistical fit (Table) and good visual fit to the cabozantinib OS KM. The regorafenib OS generated using the constant HR from the anchored MAIC is shown in Figure 19.

Figure 19: Regorafenib OS generated from anchored MAIC constant HR compared with RESORCE regorafenib OS KM



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; OS, overall survival

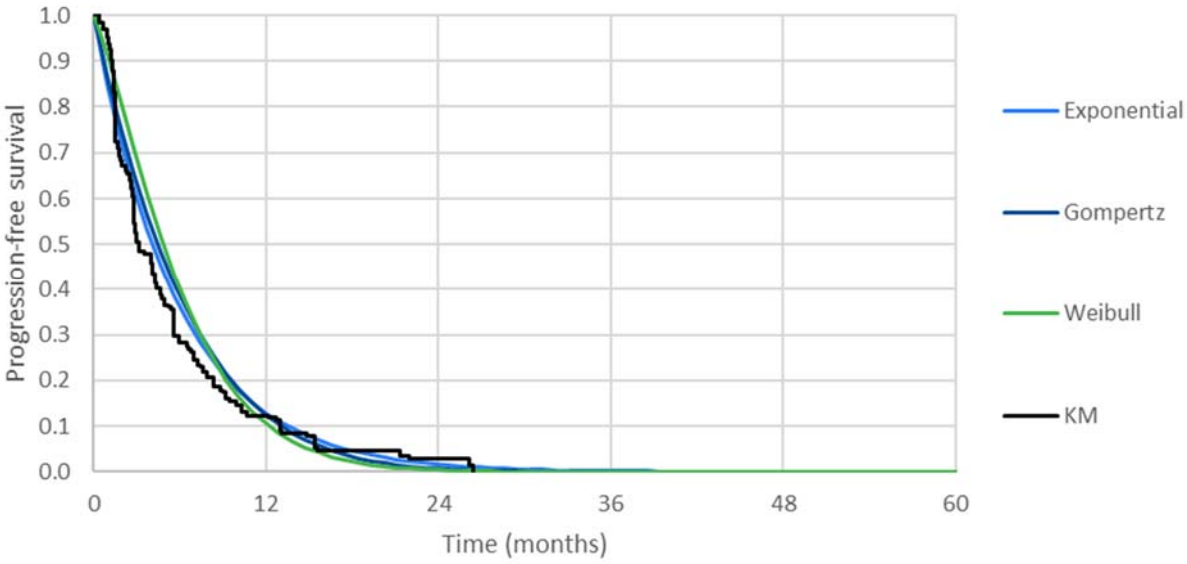
Figure 20: Base case cabozantinib and regorafenib OS from the anchored MAIC constant HR scenario



Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; OS, overall survival

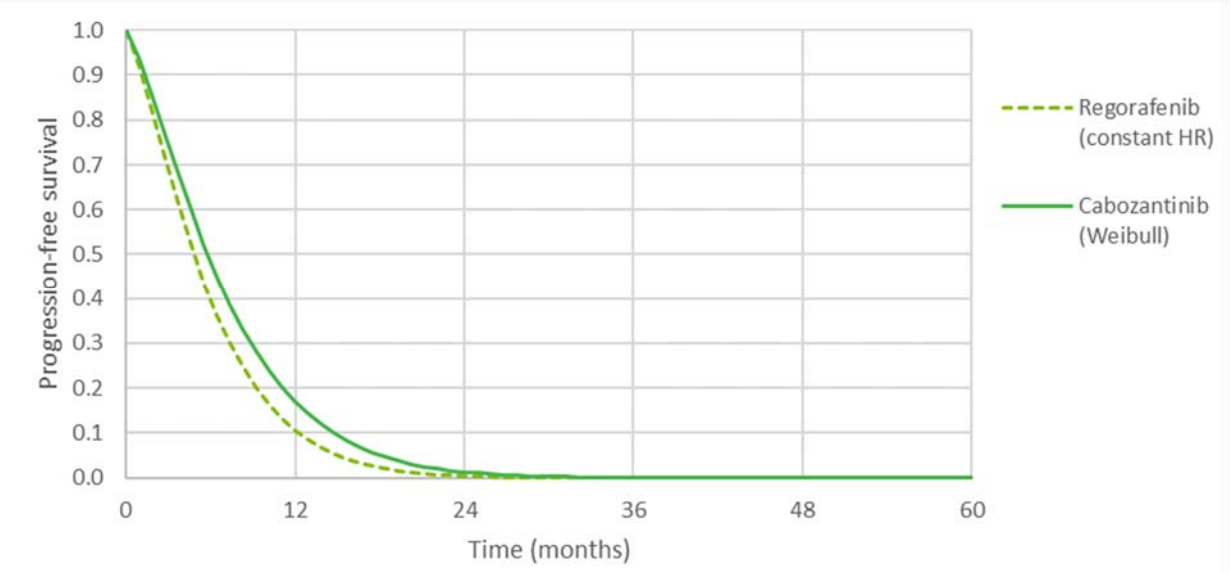
The regorafenib PFS generated using the Cox PH constant HR from the anchored MAIC is shown in Figure 21. The Weibull model was selected as the base case (Figure 22) for this scenario due to the statistical fit (Table) and good visual fit to the cabozantinib PFS KM.

Figure 21: Regorafenib PFS generated from anchored MAIC constant HR compared with RESORCE regorafenib PFS KM



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

Figure 22: Base case cabozantinib and regorafenib PFS from the anchored MAIC constant HR scenario



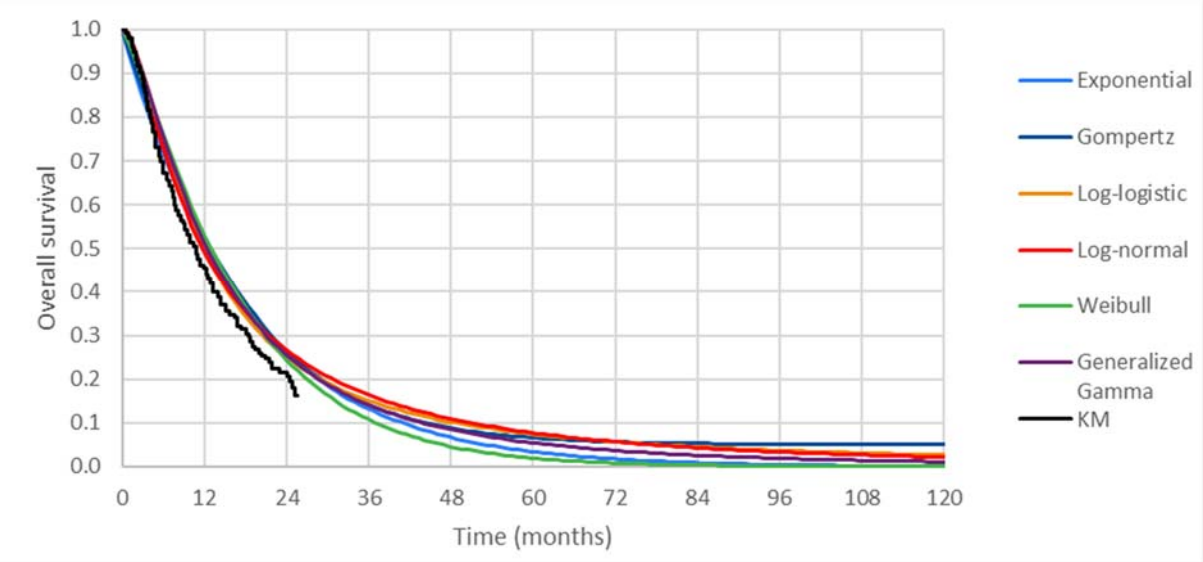
Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

Anchored MAIC time-varying HR model fitting

The regorafenib OS generated using the time-varying HR from the anchored MAIC is shown in Figure 23. The log-logistic model was selected as the base case (

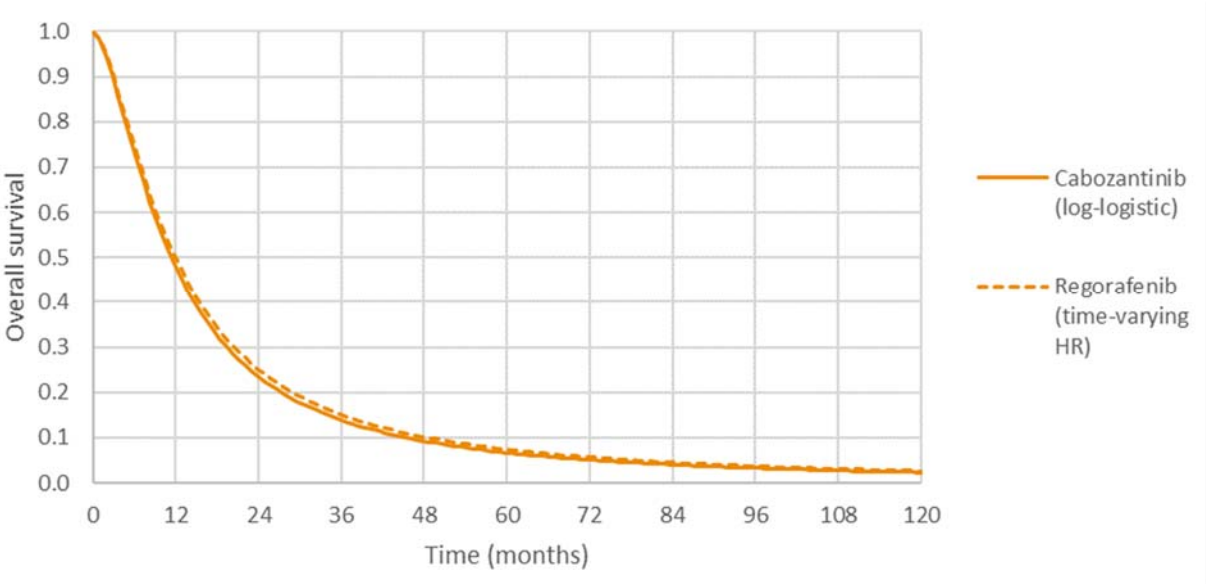
Figure 24) for this scenario as highlighted in CS, Section B.3.10.3, page 75-76 and CS, Appendix I, Section I.1.1, page 111-116. The OS for regorafenib is closer to the observed values from RESORCE than the constant HR scenarios. However, the estimated OS is still greater than the OS KM from RESORCE after approximately 6 months.

Figure 23: Regorafenib OS generated from anchored MAIC time-varying HR compared with RESORCE regorafenib OS KM



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; OS, overall survival

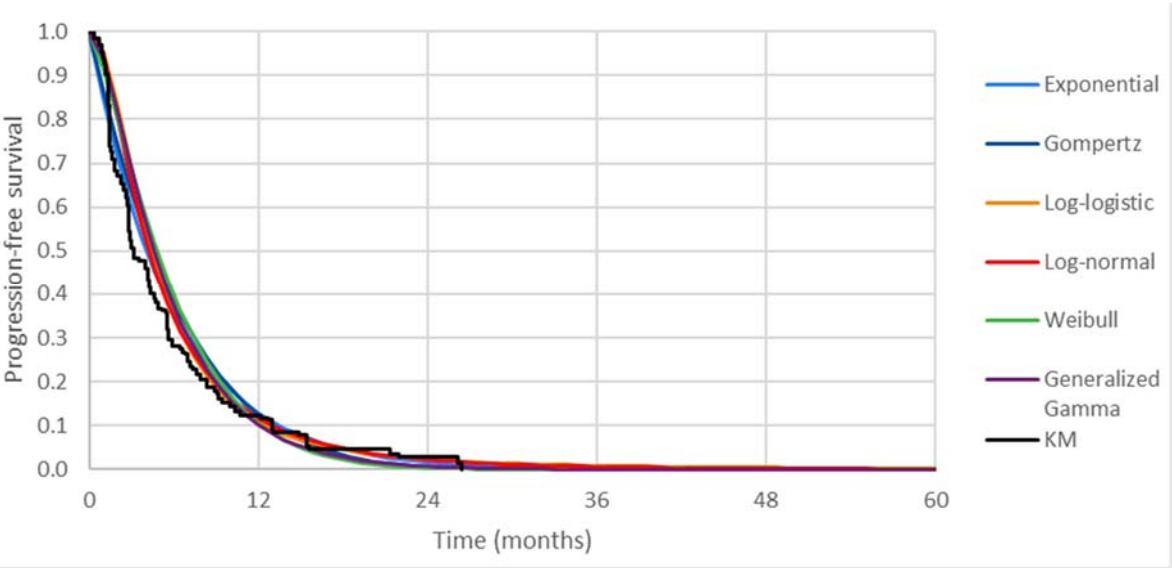
Figure 24: Base case cabozantinib and regorafenib OS from the anchored MAIC time-varying HR scenario



Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; OS, overall survival

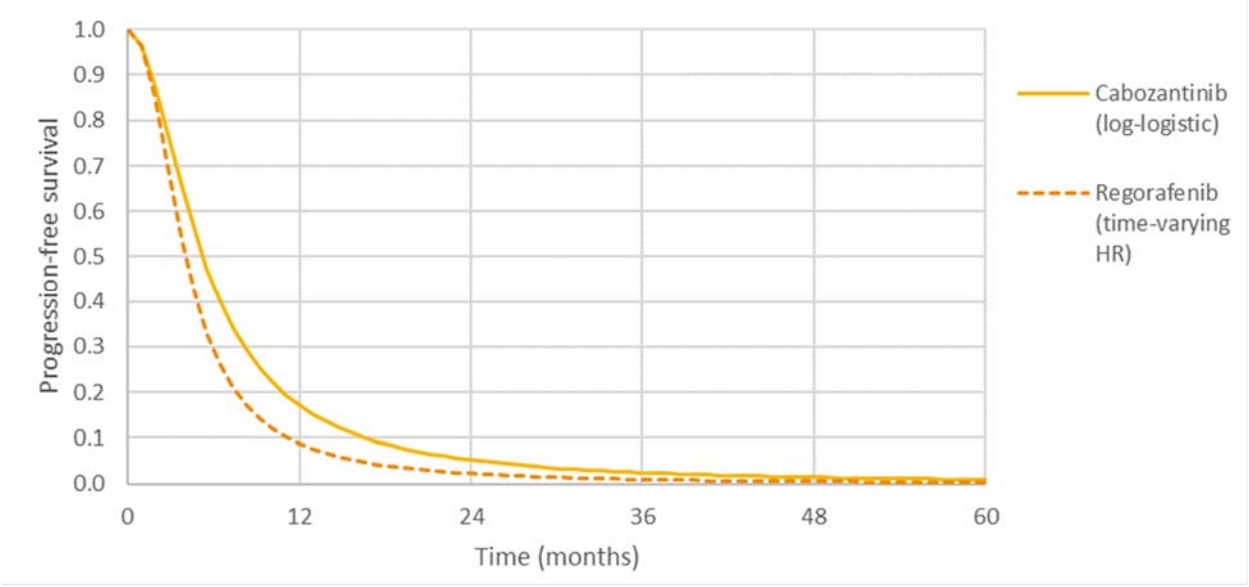
The regorafenib PFS generated using the time-varying HR from the anchored MAIC is shown in Figure 25. The log-logistic model was selected as the base case (Figure 26) for this scenario as highlighted in CS, Section B.3.10.3, page 75-76 and CS, Appendix I, Section I.1.2, page 116-121.

Figure 25: Regorafenib PFS generated from anchored MAIC time-varying HR compared with RESORCE regorafenib PFS KM



Abbreviations: HR, hazard ratio; KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

Figure 26: Base case cabozantinib and regorafenib PFS from the anchored MAIC time-varying HR scenario



Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival

Results

The deterministic incremental QALY results for the ITC scenarios are shown in Table . A 3-health state partitioned survival model with progression-free, progressed disease and death health states were used to estimate health state occupancy. The CELESTIAL trial-based utility values from Table 16 were used in the incremental QALY analysis. The deterministic results across the various MAIC approaches show a range of incremental QALYs centred around 0 QALY gain. [redacted]; both scenarios show a QALY gain from the progression-free health state and a QALY loss in the progressed disease health states, which is consistent with the HR for PFS favouring cabozantinib while favouring regorafenib for OS.

Table 18: Deterministic incremental QALY results by health state for ITC scenarios

Scenario	Progression-free incremental QALY	Progressed disease incremental QALY	Total incremental QALY
Anchored MAIC constant HR (Weibull HR base case)	[redacted]	[redacted]	[redacted]

Scenario	Progression-free incremental QALY	Progressed disease incremental QALY	Total incremental QALY
Anchored MAIC constant HR (Cox PH base case)	[redacted]	[redacted]	[redacted]
Anchored MAIC time-varying HR	[redacted]	[redacted]	[redacted]
Unanchored MAIC	[redacted]	[redacted]	[redacted]

Abbreviations: HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; PH, proportional hazard; QALY, Quality adjusted life year

A probabilistic sensitivity analysis (PSA) was performed to account for multivariate and stochastic uncertainty in the model. The implementation of PSA involved assigning specific parametric distributions and repeatedly sampling mean parameter values, 1,000 simulations were run. The distributional assumptions made for each variable were as follows:

- A lognormal distribution was used for HRs
- A beta distribution was used for utilities
- A multivariate normal distribution was used for varying survival curve parameters

The probabilistic results for each scenario are presented in Table . Figure 27 to Figure 30 visualise the incremental QALY distribution by presenting the iterations in a histogram for each scenario. The results show for each scenario a distribution of positive and negative incremental QALYs with collectively the most frequent iterations near the 0 QALY gain point estimate, demonstrating no meaningful difference in QALYs between cabozantinib and regorafenib in a pure second line HCC population previously treated with sorafenib irrespective of tolerability.

Table 19: Probabilistic incremental QALY results by health state for ITC scenarios

Scenario	Total incremental QALY (mean, [SE])
Anchored MAIC constant HR (Weibull HR base case)	[redacted]
Anchored MAIC constant HR (Cox PH)	[redacted]
Anchored MAIC time-varying HR	[redacted]
Unanchored MAIC	[redacted]

Abbreviations: HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; PH, proportional hazards; QALY, Quality adjusted life year; SE, standard error

Figure 27: Histogram of incremental QALYs from PSA for the anchored MAIC constant HR (Weibull HR) scenario

Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PSA, probability sensitivity analysis; QALY, quality adjusted life-year

Figure 28: Histogram of incremental QALYs from PSA for the anchored MAIC constant HR (Cox PH) scenario

Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PH, proportional hazards; PSA, probability sensitivity analysis; QALY, quality adjusted life-year

Figure 29: Histogram of incremental QALYs from PSA for the anchored MAIC time-varying HR scenario

Abbreviations: HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PSA, probability sensitivity analysis; QALY, quality adjusted life-year

Figure 30: Histogram of incremental QALYs from PSA for the unanchored MAIC scenario

Abbreviations: MAIC, matching-adjusted indirect comparison; PSA, probability sensitivity analysis; QALY, quality adjusted life-year

References

Abou-Alfa GK, Meyer T, Cheng AL, El-Khoueiry AB, Rimassa L, Ryoo BY, Cicin I, Merle P, Chen Y, Park JW, Blanc JF, Bolondi L, Klümpen HJ, Chan SL, Zagonel V, Pressiani T, Ryu MH, Venook AP, Hessel C, Borgman-Hagey AE, Schwab G, Kelley RK. Cabozantinib in Patients with Advanced and Progressing Hepatocellular Carcinoma. N Engl J Med. 2018 Jul 5;379(1):54-63. doi: 10.1056/NEJMoa1717002. PMID: 29972759; PMCID: PMC7523244.

Bayer Plc, 2022. Summary of product characteristics: Stivarga 40 mg film-coated tablets. https://www.medicines.org.uk/emc/product/1263. Accessed 16 March, 2022.

Bayer Plc, 2022. Summary of product characteristics: Stivarga 40 mg film-coated tablets. https://www.medicines.org.uk/emc/product/1263. Accessed 16 March, 2022.

Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, Pracht M, Yokosuka O, Rosmorduc O, Breder V, Gerolami R, Masi G, Ross PJ, Song T, Bronowicki JP, Ollivier-Hourmand I, Kudo M, Cheng AL, Llovet JM, Finn RS, LeBerre MA, Baumhauer A, Meinhardt G, Han G; RESORCE Investigators. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2017 Jan 7;389(10064):56-66. doi: 10.1016/S0140-6736(16)32453-9. Epub 2016 Dec 6. Erratum in: Lancet. 2017 Jan 7;389(10064):36. PMID: 27932229.

Casadei-Gardini A, Rimassa L, Rimini M, Yoo C, Ryoo BY, Lonardi S, Masi G, Kim HD, Vivaldi C, Ryu MH, Rizzato MD, Salani F, Bang Y, Pellino A, Catanese S, Burgio V, Cascinu S, Cucchetti A. Regorafenib versus cabozantinb as second-line treatment after sorafenib for unresectable hepatocellular carcinoma: matching-adjusted indirect comparison analysis. J Cancer Res Clin Oncol. 2021 Dec;147(12):3665-3671. doi: 10.1007/s00432-021-03602-w. Epub 2021 Mar 20. PMID: 33745079.

European Medicines Agency, 2021 Summary of Product Characteristics – Stivarga. https://www.ema.europa.eu/en/documents/product-information/stivarga-epar-product-information_en.pdf. Accessed 16 March 2022.

Finn RS, Merle P, Granito A, Huang YH, Bodoky G, Pracht M, Yokosuka O, Rosmorduc O, Gerolami R, Caparello C, Cabrera R, Chang C, Sun W, LeBerre MA, Baumhauer A, Meinhardt G, Bruix J. Outcomes of sequential treatment with sorafenib followed by regorafenib for HCC: Additional analyses from the phase III RESORCE trial. J Hepatol. 2018 Aug;69(2):353-358. doi: 10.1016/j.jhep.2018.04.010. Epub 2018 Apr 26. PMID: 29704513.

Firtina Karagonlar Z, Koc D, Iscan E, Erdal E, Atabey N. Elevated hepatocyte growth factor expression as an autocrine c-Met activation mechanism in acquired resistance to sorafenib in hepatocellular carcinoma cells. Cancer Sci 2016; 107: 407-16.

Liu J, Wang K, Yan Z, et al. Axl expression stratifies patients with poor prognosis after hepatectomy for hepatocellular carcinoma. PLoS One 2016; 11(5): e0154767.

Rimassa L, Abbadessa G, Personeni N, et al. Tumor and circulating biomarkers in patients with second-line hepatocellular carcinoma from the randomized phase II study with tivantinib. Oncotarget 2016; 7: 72622-33.

Rimassa L, Assenat E, Peck-Radosavljevic M, et al. Tivantinib for second-line treatment of MET-high, advanced hepatocellular carcinoma (METIV-HCC): a final analysis of a phase 3, randomised, placebo-controlled study. Lancet Oncol 2018; 19: 682-93.

Santoni M, Iacovelli R, Colonna V, et al. Antitumor effects of the multi-target tyrosine kinase inhibitor cabozantinib: a comprehensive review of the preclinical evidence. Expert Review of Anticancer Therapy, 2021; 21(9): 1029-1054.

Ueki T, Fujimoto J, Suzuki T, Yamamoto H, Okamoto E. Expression of hepatocyte growth factor and its receptor, the c-met proto-oncogene, in hepatocellular carcinoma. Hepatology 1997; 25: 619-23.

WHOCC - ATC/DDD Index. 2021 https://www.whocc.no/atc_ddd_index/?code=L01EX&showdescription=yes. Accessed 16 March 2022.

Xiang Q, Chen W, Ren M, et al. Cabozantinib suppresses tumor growth and metastasis in hepatocellular carcinoma by a dual blockade of VEGFR2 and MET. Clin Cancer Res 2014; 20: 2959-70.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Vanessa Hebditch**

2. Name of organisation	British Liver Trust
3. Job title or position	Director of Communications and Policy
4a. Brief description of the organisation (including who funds it). How many members does it have?	The British Liver Trust is the UK's leading liver health charity working to improve liver health for all and supporting all adults affected by liver disease or liver cancer. We are funded by voluntary donations including community and event fundraising, individual donors, gifts in wills, corporate supporters and trust and foundation grants. We operate throughout the UK and reach over a million people each year. Our website has over 1.5 million unique visitors each year; our online forum has over 29,000 active members, our nurse-led Helpline handles between 400 and 500 enquiries a month, our regular newsletter goes to c17,000 people with liver disease and liver cancer, we run around 250 support groups each year (currently virtual but moving to a mix of virtual and face to face post Covid); we expect to visit around 40 locations per annum with our Love Your Liver Roadshow which raises awareness of the risk factors of liver disease, we connect with around 20,000 people via social media.
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? If so, please state the name of manufacturer, amount, and purpose of funding.	The British Liver Trust received an educational grant of £10,600 from Ipsen for the production of patient materials for liver cancer patients and in support of raising awareness and launching these materials during Liver Cancer Awareness Month in October 2021. Ipsen had no control or influence over the content or promotion of these materials (which were co-produced by patients, carers and clinical experts).
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No.

5. How did you gather information about the experiences of patients and carers to include in your submission?	Information in this submission is collated from a variety of British Liver Trust sources and activities, including: - Direct feedback and intelligence from patients and carers who contact the British Liver Trust specialist nurse helpline - Direct feedback and intelligence from patients and carers who attend British Liver Trust patient support groups - Feedback through focus groups of people living with liver cancer and those who care for them - Literature reviews - Results of patient surveys including a survey of over 2000 people living with liver disease and liver cancer and a separate survey of 127 patients living with hepatocellular carcinoma (HCC) - Extended one to one interviews with 5 patients with HCC - Responses through website and social media channels - Feedback via threads on our online patient forum (over 29,000 members) - Feedback from our Patient Advisory Group members - Intelligence and information from our Clinical Advisory Group regarding issues that they hear from patients
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Primary liver cancer (HCC) is complex, varied and fluctuates, meaning that no one person's experience is the same as another. Many patients (approx. 80%) also have underlying liver cirrhosis, which not only makes treatment difficult but also means that they may have other complications. They live with uncertainty, hopelessness and often stigma and isolation due to the image of liver cancer. Primary liver cancer in adults has a poor outlook because it tends to be diagnosed late (only 10% of people are diagnosed in the early stages, when surgery can help). The five-year survival rate is only 12-15%. For people where surgery is not an option, the prognosis is particularly poor, and it is rare for people to live more than three years. The lack of other chemotherapeutic drugs particularly affects this group as well as those awaiting a transplant.

	Because patients with advanced HCC have such a poor prognosis and there are so few treatment options they are usually completely devastated. Patients are often relatively young and are completely shell shocked. Patients also report feeling extremely unwell, very tired and weak. Some quotes from a focus group and one to one interviews with patients and carers include: "Emotionally it was tough. I felt like I couldn't cope and it all just caught up with me. I felt like every time I put my head up above water I got shot down." "Immediately after diagnosis I was shell shocked. I took my house in order, made my will. But I made changes to things. Death was imminent in my mind. Having a transplant makes me realise how lucky I am but I wish there had been another option. Liver disease doesn't seem to get the attention of other cancers." "We were just devastated. My husband was prescribed medication and underwent a radiofrequency ablation procedure. He was extremely tired and in pain. He was put on the waiting list, then he had to be taken off the list as the cancer had grown whilst waiting. He was 42 years old, had never drunk in his life and we were told he would die in about six weeks. The rug was completely taken from under my feet … my whole life crumbled and ten years on I am still in pain." Relatives have described the condition as "brutal - the worst possible way to go". When patients are diagnosed with HCC, they often experience depression from the poor prognosis and a range of symptoms including severe pain that cannot be treated without worsening their liver condition. Other severe symptoms include ascites, fluid in the abdomen that can press on the stomach making it difficult to eat and even to breathe. Hepatic encephalopathy can make everyday functions including conversation, writing and staying awake difficult. Only a very few patients are offered curative treatment, and even then, many live with the uncertainty about whether they will receive a liver transplant before the tumour spreads, or whether they will die as a complication of surgery (liver resection has a relatively high mortality rate). Patients with HCC are often many years younger than those with other cancers, and extra time is of particular importance to people who may have young families and working lives to put in order before death.

	Buying extra time for such patients not only can positively impact those individuals, but can also have a huge positive impact on families and the wider community, with unquantifiable downstream benefits that can bring. There is also wide variation of care across England and Wales with patients experiencing different standards of care depending on where they live. Our survey of 127 patients with HCC revealed • 90% wanted more information after they left their first appointment at the hospital • 1 in 5 were not happy with the information they received about treatment options • Nearly half of respondents (45%) said they asked their doctor for other treatments they had researched that were not initially offered • More than one in ten (13%) said their liver cancer diagnosis began with a trip to A&E because of symptoms • One in five (21%) said it took more than six months to get a liver cancer diagnosis after their first visit to the GP • 44% respondents said they have experienced delays in accessing care since their diagnosis. • Half of patients (49%) didn't have treatment at their local hospital and over half (55%) said they travelled 20 miles or more for their treatment • 44% respondents said they have experienced delays in accessing care since their diagnosis. • Half (51%) said the COVID-19 pandemic has affected their care

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Patients are really shocked when they realise the lack of treatment options. When there is no option for surgical treatments, minimally invasive therapies such as TACE or SIRT or liver transplant the current only life extending treatment options for patients with advanced liver cancer is sorafenib (Nexevar). Patients report side effects and for some people these are severe. Once sorafenib stops working, the only option is currently Regorafenib (Stirvaga). Once these options are exhausted the only option is palliative care. HCC patients are disadvantaged purely because they have a disease which does not have an extensive number of treatments available. For example, in many other cancers, there are several life-extending chemotherapy treatments available, and it may be appropriate to consider whether new medicines are effective. This is not the case in liver cancer.
8. Is there an unmet need for patients with this condition?	Yes. HCC has a poor survival prognosis. It is a debilitating condition with many distressing symptoms. These patients have limited treatment options.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	The British Liver Trust has not spoken to any patients who have received this treatment in clinical trials. However, patients are desperate for any new treatments and were encouraged by the data that has been published in peer review journals. They saw it as a much needed and welcome additional treatment option for use in adults with HCC. Improving quality of life and even small extensions to length of life are of considerable importance to this patient group.

Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Patients understood that there could be side effects but believed that these would be tolerable and acceptable and that any adverse events would be manageable. Patients spoke about being in "last chance saloon" and willing to put up with this for an extended life. Many patients with HCC are relatively young and so they cherished the possibility of having last moments such as "seeing their grandchild" "spending special time with family". An extension of life was also seen as an opportunity to put "their house in order". Many patients reported having side effects with sorafenib and talked about how much they would appreciate a further treatment option. Having read the literature reviews they believed that the main side effects of cabozantinib (cabapalmar–plantar erythrodysesthesia, hypertension, fatigue, and diarrhea) could be less severe than those they had been experiencing on sorafenib.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	

Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Liver disease and liver cancer disproportionally affects the poorest in society. Many patients with liver cancer come from disadvantaged backgrounds and have complex lives.
Other issues	
13. Are there any other issues that you would like the committee to consider?	
Key messages	

14. In up to 5 bullet points, please summarise the key messages of your submission:

- A diagnosis of liver cancer is devastating and the prognosis is very poor (average 5 year survival of 13 years)
- There are very few treatment options currently available
- Any new treatment that may prolonged their life and provided them with a real chance of survival is desperately needed for these patients
- Patients with HCC are often many years younger than those with other cancers, and extra time is of particular importance to people who may have young families and working lives to put in order before death

NICE National Institute for Health and Care Excellence

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

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

Professional organisation submission

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**British Association for the Study of the Liver (BASL) / HCC UK**

3. Job title or position	**Consultant medical oncologist**
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	The British Association for the Study of the Liver is the National Association for hepatology. BASL is composed of interested individuals from clinical medicine, clinical and basic research and allied professions. BASL is funded through membership fees and organising and hosting an annual meeting and educational events. HCC-UK is a national cross-specialty group of clinicians with an interest in hepatocellular carcinoma (HCC) and a special interest group of BASL.
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.]	Yes – BAYER (comparator) BASL received £550.00 in sponsorship funding towards an annual meeting of HCC-UK that took place in March 2021.

If so, please state the name of manufacturer, amount, and purpose of funding.	
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**NO**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	Prolong overall survival Prolong progression free survival Maintain quality of life
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by	A clinical significant treatment response would be to improve median overall survival by ≥3months.

x cm, or a reduction in disease activity by a certain amount.)	
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, there are limited treatment options available for patient with advanced HCC who have progressed on sorafenib (regorafenib or BSC) or who are intolerant of sorafenib (lenvatinib or BSC). For patients who receive lenvatinib for advanced HCC there is no available active therapy on disease progression (regorafenib not funded for these patients).
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Patient with advanced HCC who have progressed on sorafenib are currently treated with either regorafenib or BSC. Patients with advanced HCC who are intolerant of sorafenib are currently treated with lenvatinib or BSC.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	BCLC guidelines EASL guidelines
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please	The pathway of care is well defined, however patients may either receive sorafenib or lenvatinib if unsuitable for atezo/bev (or after progression on atezo/bev. If patients receive lenvatinib they do not have any further active therapy available by NHS funding (regorafenib only funded after sorafenib).

state if your experience is from outside England.)	
• What impact would the technology have on the current pathway of care?	Funding for cabozantinib would offer patients who progress following treatment with lenvatinib a further option for active therapy. In addition funding for lenvatinib would offer a further line of active therapy for patients who progress after sorafenib/regorafenib
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes
• How does healthcare resource use differ between the technology and current care?	No difference
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care oncology clinics
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None

11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes – clinically meaningful benefit compared to BSC
• Do you expect the technology to increase length of life more than current care?	Yes – 3.3month increase in median survival compared to BSC/placebo
• Do you expect the technology to increase health-related quality of life more than current care?	Unable to comment. No relevant data published in the Celestial trial.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No predictive biomarkers available.
The use of the technology	

13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	No difference from current care. No practical implications
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Treatment would stop on development of radiological progression or intolerable toxicity.
15. Do you consider that the use of the technology will result in any substantial health-	No

related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes, the technology is innovative, and would lead to a statistically significant and clinically meaningful increase in median overall survival.
• Is the technology a 'step-change' in the management of the condition?	The technology is not a 'step-change' in management as there are other multi-kinase inhibitor drugs already used for this condition.
• Does the use of the technology address any particular unmet need of the patient population?	Funding for cabozantinib would offer patients who progress following treatment with lenvatinib a further option for active therapy. In addition funding for lenvatinib would offer a further line of active therapy for patients who progress after sorafenib/regorafenib

17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Side effects may potentially negatively impact on quality of life, however improved disease control is likely to positively impact on quality of life.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	The clinical trial (Celestial) doesn't reflect current UK clinical practice as the majority of patients will now receive atez/bev first line rather than sorafenib.
• If not, how could the results be extrapolated to the UK setting?	However it is reasonable to extrapolate the Celestial trial results to patients who received atezo/bev first line and sorafenib second-line. In addition some patients will not be suitable for atezo/bev and will hence receive sorafenib or lenvatinib as first-line therapy.
• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival - reported by the Celestial trial
• If surrogate outcome measures were used, do they adequately predict	None used.

long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
20. Are you aware of any new evidence for the comparator treatment since the publication of NICE technology appraisal guidance TA555?	No
21. How do data on real-world experience compare with the trial data?	None available.

Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	
Key messages	
23. In up to 5 bullet points, please summarise the key messages of your submission. • Cabozanitnib offers patients with advanced HCC a meaningful improvement in overall survival • Side effects reported in the Celestial trial were in line with other similar drugs and manageable • There are limited treatment options for patients with previously treated HCC • •	

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

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

Professional organisation submission

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**NCRI-ACP-RCP-RCR**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	**NCRI-ACP-RCP-RCR**
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.]	**No**

If so, please state the name of manufacturer, amount, and purpose of funding.	
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by	

x cm, or a reduction in disease activity by a certain amount.)	
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Currently, management of advanced HCC relies on the use of atezolizumab and bevacizumab, sorafenib or Lenvatinib in the first-line setting. After progression to sorafenib, and only if sorafenib was well-tolerated, regorafenib is a subsequent line of treatment available. However, no alternatives to regorafenib are available for the significant proportion of patients who did not tolerate sorafenib well. No significant geographical variations exist. I don't think that relevant differences between professionals exist regarding the benefit that cabozantinib would represent for patients with advanced HCC. There is no alternative available for patients who did not tolerate sorafenib. Regorafenib could be an alternative for patient who did tolerate sorafenib well.
• 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?	
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Cabozantinib is currently not used, since it is not available.
• How does healthcare resource use differ between the technology and current care?	
• In what clinical setting should the technology be used? (For example,	Resources that are currently available for delivering treatment options such as sorafenib, Lenvatinib and regorafenib would be used. I do not foresee any need for additional resources to be required

primary or secondary care, specialist clinics.)	
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	NHS staff already trained for delivery and management of sorafenib, Lenvatinib and regorafenib would have sufficient training for the management of cabozantinib. I would not expect additional training to be required.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Patient would have an additional option of treatment in the scenario of a highly unmet need
• 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?	

12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	

14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it	

improve the way that current need is met?	
• 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?	
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	
Sources of evidence	
18. 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?	
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	
20. Are you aware of any new evidence for the comparator	

treatment since the publication of NICE technology appraisal guidance TA555?	
21. How do data on real-world experience compare with the trial data?	
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	
22b. Consider whether these issues are different from issues with current care and why.	
Key messages	

23. In up to 5 bullet points, please summarise the key messages of your submission.

-
-
-
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

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

Clinical expert statement

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Thank you for agreeing to comment on the evidence review group (ERG) report for this appraisal, 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 ERG report and stakeholder responses are used by the appraisal committee to help it make decisions at the appraisal 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 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.

NICE National Institute for Health and Care Excellence

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 replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **Wednesday 23 March**. 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 advanced hepatocellular carcinoma and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Tim Meyer
2. Name of organisation	University College London
3. Job title or position	
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 advanced hepatocellular carcinoma? ☐ A specialist in the clinical evidence base for advanced hepatocellular carcinoma 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 advanced hepatocellular carcinoma?	To improve survival by delaying disease progression.

(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)	To provide a statistically significant hazard ratio for death of at least 0.8 or in other words to reduce the risk of death by 20% compared to the standard of care.
10. In your view, is there an unmet need for patients and healthcare professionals in advanced hepatocellular carcinoma?	Hepatocellular carcinoma is the third leading cause of cancer death worldwide and has one of the lowest five-year survivals of all cancers at around 8%. The majority of patients are not suitable for curative intervention and are treated with locoregional or systemic therapy. Improving the efficacy of systemic therapy is critical for delivery of better outcomes in advanced disease and remains a significant unmet need.
11. How is advanced hepatocellular carcinoma 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?	The following clinical guidelines are in use: 1. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma J Hepatology 2018. 2. Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2018. 3. ILCA Systemic therapy guidance https://ilca-online.org/wp-content/uploads/2020/06/Systemic-therapy-guidelines-V1.2.pdf . All define a similar therapeutic algorithm for systemic therapy. There are three evidence based first line therapies available for advanced HCC. The most effective is the combination of Atezolizumab and Bevacizumab (AB) which is associated with an objective response rate of 27% and median overall survival of 19 months. For patients unable to receive AB, sorafenib has been shown to improve survival compared with placebo and lenvatinib has been shown to be non-inferior to sorafenib. Either are recommended as equivalent first-line options as an alternative to AB or in those who progress on or do not tolerate AB. For patients who have received sorafenib, there are positive placebo controlled trials supporting second line use of 1. regorafenib in those that tolerated sorafenib 2. Cabozantinib and 3.ramucirumab in those with AFP ≥ 400ng/ml. In the trial of cabozantinib, 28% had received 2 prior lines of therapy. Currently in the UK, only regorafenib is approved for use following sorafenib. The approval of cabozantinib would provide an alternative to regorafenib with a broader applicability in that the registrational trial did not mandate tolerance of

	sorafenib. This is important since sorafenib is often poorly tolerated and around 20% patients discontinue treatment due to poor tolerance.
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)	Cabozantinib is recommended as a second line therapy following sorafenib. It will be prescribed in specialist services for the management of HCC which have the requisite multidisciplinary team capable of managing both the cancer and the chronic liver disease which is present in around 90% of those with HCC. The optimal setting is a joint clinic staffed by both oncologists and hepatologists. These clinics are familiar with administration of this class of drug for patients with HCC and no additional training or infrastructure should 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 reduce disease progression more than current care? • What proportion of patients are expected to be progression-free at 2 years and 4 years? • Do you expect the technology to increase health-related quality of life more than current care?	Cabozantinib improves survival compared with placebo and has a broader applicability than regorafenib. If approved, it is likely to become the drug of choice as second line therapy following sorafenib. Compared with placebo progression free survival was improved from 1.9 to 5.2 months and overall survival from 8.0 to 10.2 months. Cabozantinib has not been compared with regorafenib in a clinical trial. The proportion of patients expected to be progression free at 2 or 4 years is less than 1%. By delaying progression, disease related symptoms will be delayed. A peer reviewed publication by Freemantle N et al has been accepted by European Journal of Cancer and will be online shortly. In this publication, quality of life has been formally assessed using validate tool (EQ-5D-5L) with the context of the placebo controlled Celestial trial, and significant improvements in mean QALY were identified in favour of cabozantinib.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	The Celestial trial was conducted in patients with well preserved liver function (Child Pugh A disease) and good performance status (ECOG PS 0 or 1). Published data (Kelley et al Brit J Cancer 2020) show that the absolute benefit is less in those with impaired liver function which is a consistent finding in this

	disease. Confining treatment to those with Child Pugh A liver disease would seem appropriate.
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)	Health care professionals are familiar with use of tyrosine kinase inhibitors such as cabozantinib for the treatment of HCC. They are oral drugs administered daily with routine outpatient monitoring.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Patients will usually be monitored with CT or MRI imaging every 2-3 months and treatment will be stopped if the patient chooses or if there is loss of clinical benefit or if the disease progresses radiologically.
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?	The technology is likely to be better tolerated than regorafenib in those that are intolerant of sorafenib and therefore this drug has broader applicability.

• 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? • How are adverse events of grade 3 or more typically managed? • Are there substantial costs associated with the management of adverse effects?	The side effect profile of cabozantinib is typical for this class of drugs and is managed in the outpatient setting with supportive medication or dose reduction. The most common side effects requiring supportive medication are diarrhoea, palmer planter erythrodysesthesia and hypertension which occur at grade three level in 10%, 17% and 16% respectively in the Celestial trial. These events were reported irrespective of causality and 34% patients in the placebo group also recorded grade 3 events. In the celestial trial, Grade 4 events occurred in 10% of those on cabozantinib and 3% on placebo but there is no consistent event. Since the rate of need for supportive medication is low and the cost of those supportive drugs is also low and all toxicities are managed as an outpatient, the costs associated with managing side effects is not substantial.
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?	Although the Celestial trial was global, around 50% trial recruitment was from Europe and the patient population is representative of that of the UK The most important outcomes are overall survival, progression free survival, response rate, toxicity and health related quality of life all of which were measure in the Celestial trial. I am not aware of any new adverse events arising in post marketing studies
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	Casadei-Gardini A et al 2021 J Cancer Res Clin Oncol. Regorafenib versus cabozantinb as second-line treatment after sorafenib for unresectable hepatocellular carcinoma: matching-adjusted indirect comparison analysis

	Conclusion: Our results confirmed no differences between regorafenib and cabozantinib in terms of OS. However, in earlier progressors on prior sorafenib a larger benefit might be expected from cabozantinib treatment.
22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance TA555?	No
23. How do data on real-world experience compare with the trial data?	There is very little published
24. NICE considers whether there are any equalities issues at each stage of an appraisal. 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. Please state if you think this appraisal 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	No

• 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.	
25. Have there been substantial changes to the treatment pathway since the regorafenib appraisal (TA555) that might impact the relevance of the comparator's appraisal?	With the approval of atezolizumab and bevacizumab as first line therapy, less patients are receiving sorafenib and the patient population for which cabozantinib may be considered has reduced as a consequence. But there are no other comparators other than regorafenib in the UK and ramucirumab outside the UK
26. How often are patients offered regorafenib in this population, compared to other treatment options. I.e., what is the current market share of regorafenib in this indication?	Regorafenib is the only drug approved in the UK for this indication.
27. How is regorafenib being used in clinical practice?	Second line after first line sorafenib or third line after second line sorafenib
28. What proportion of people are being treated with regorafenib in the second and third-line setting?	I would estimate that the majority will be third line since less than 10% will have first line sorafenib.
30. Is the CELESTIAL trial generalisable to UK patients with advanced hepatocellular carcinoma?	Yes
31. What is the tolerability of cabozantinib compared with regorafenib?	There has been no direct trial comparing cabozantinib with regorafenib and they are similar class of drug. Comparative toxicity was evaluated in Kelley RK et al Adv Ther PMID 32424805 which compared the data from the Celestial and Resorce trial but this was potentially biased by the fact that Resorce preselected patients tolerant to sorafenib and therefore excluded those likely to get side effects from tyrosine kinase inhibitors.

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Cabozantinib is a recommended option after sorafenib in all international guidelines
Cabozantinib has demonstrated clinically significant improvement in survival, progression free survival and QOL compared to placebo
The toxicity profile is well defined and side effects can be managed as an outpatient with low cost supportive medication when needed
Cabozantinib has broader applicability than regorafenib which was only evaluated in sorafenib tolerant population
The proportion of patients who will be eligible has reduced with the introduction of AtezoBev as first line standard of care.

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

Clinical expert statement

Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]

Thank you for agreeing to comment on the evidence review group (ERG) report for this appraisal, 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 ERG report and stakeholder responses are used by the appraisal committee to help it make decisions at the appraisal 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 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.

NICE National Institute for Health and Care Excellence

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 replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **Wednesday 23 March**. 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.

NICE National Institute for Health and Care Excellence

Part 1: Treating advanced hepatocellular carcinoma 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 Richard Hubner
2. Name of organisation	The Christie NHS Foundation Trust
3. Job title or position	Consultant in Medical Oncology
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 advanced hepatocellular carcinoma? ☒ A specialist in the clinical evidence base for advanced hepatocellular carcinoma 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, I wrote the organsation submission (BASL/HCC-UK)
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	
8. What is the main aim of treatment for advanced hepatocellular carcinoma?	

(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)	
10. In your view, is there an unmet need for patients and healthcare professionals in advanced hepatocellular carcinoma?	
11. How is advanced hepatocellular carcinoma 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?	
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)	

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 reduce disease progression more than current care? • What proportion of patients are expected to be progression-free at 2 years and 4 years? • Do you expect the technology to increase health-related quality of life more than current care?	
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	
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)	
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any 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	
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?	
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? • How are adverse events of grade 3 or more typically managed? • Are there substantial costs associated with the management 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? • 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?	

21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	
22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance TA555?	
23. How do data on real-world experience compare with the trial data?	
24. NICE considers whether there are any equalities issues at each stage of an appraisal. 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. Please state if you think this appraisal 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.	
25. Have there been substantial changes to the treatment pathway since the regorafenib appraisal (TA555) that might impact the relevance of the comparator's appraisal?	
26. How often are patients offered regorafenib in this population, compared to other treatment options. I.e., what is the current market share of regorafenib in this indication?	
27. How is regorafenib being used in clinical practice?	
28. What proportion of people are being treated with regorafenib in the second and third-line setting?	
30. Is the CELESTIAL trial generalisable to UK patients with advanced hepatocellular carcinoma?	
31. What is the tolerability of cabozantinib compared with regorafenib?	

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

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.

Cabozantinib for previously treated advanced hepatocellular carcinoma (review of TA582). A Technology Appraisal

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 Rebecca Harvey, Statistician, Cabourn Statistics Ltd., France Katy Cooper, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK
Correspondence Author	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK
Date completed	21st April 2022

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR134849.

Declared competing interests of the authors

None of the authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Dr Rohini Sharma and Dr Sue Darby for providing clinical advice over the course of the appraisal. We would also like to thank 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, Harvey R, Cooper K, Wong R. Cabozantinib for previously treated advanced hepatocellular carcinoma (review of TA582). A Technology Appraisal. School of Health and Related Research (ScHARR), 2022.

Contributions of authors

Ruth Wong critiqued the company's search strategy. Katy Cooper summarised and critiqued the clinical effectiveness data reported within the company's submission. Rebecca Harvey critiqued the statistical aspects of the submission. Paul Tappenden critiqued the company's health economic analyses. All authors were involved in drafting and commenting on the final report.

CONTENTS

List of Tables

List of Figures

List of Boxes

Abbreviations

AE	Adverse event
AFP	Alpha-fetoprotein
AIC	Akaike Information Criterion
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AUC	Area under the curve
BCLC	Barcelona Clinic Liver Cancer
BIC	Bayesian Information Criterion
BNF	British National Formulary
BSC	Best supportive care
CDF	Cancer Drugs Fund
CI	Confidence interval
cPAS	Comparator Patient Access Scheme
CR	Complete response
CRUK	Cancer Research UK
CS	Company's submission
CSR	Clinical Study Report
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EPAR	European Public Assessment Report
EQ-5D	Euroqol 5-Dimensions
EQ-5D-5L	Euroqol 5-Dimensions 5-Level
ERG	Evidence Review Group
ESS	Effective sample size
FTA	Fast-Track Appraisal
HCC	Hepatocellular carcinoma
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient data
ITC	Indirect treatment comparisons
ITT	Intention-to-treat
LOR	Log odds ratio
MAIC	Matching-adjusted indirect comparison
mg	Milligram
mRECIST	Modified Response Evaluation Criteria In Solid Tumours
N	Number
NA	Not applicable
NASH	Non-alcoholic steatohepatitis
NHS	National Health Service
NHSE	NHS England
NICE	National Institute for Health and Care Excellence
NR	Not reported
OR	Odds ratio
ORR	Objective response rate
OS	Overall survival
PAS	Patient Access Scheme
PD	Progressive disease
PFS	Progression-free survival
PH	Proportional hazards
PPES	Palmar-plantar erythrodysaesthesia syndrome
PR	Partial response
PS	Performance status

PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria In Solid Tumours
SA	Sensitivity analysis
SAE	Serious adverse event
SmPC	Summary of Product Characteristics
TA	Technology Appraisal
TACE	Transarterial chemoembolisation
TKI	Tyrosine kinase inhibitor
ToT	Time on treatment
TTD	Time to treatment discontinuation
TTP	Time to progression
VAS	Visual analogue scale

1. SUMMARY OF THE ERG'S VIEW OF THE COMPANY'S FTA CASE

The company is seeking a positive NICE recommendation for cabozantinib in the same indication as the existing NICE recommendation for regorafenib (in TA555), that is, for the treatment of advanced unresectable hepatocellular carcinoma (HCC) in adults who have had sorafenib, only if they have Child–Pugh grade A liver impairment and an Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 or 1. This intended positioning means that the target population for cabozantinib is narrower than the patient population defined in the final scope issued by the National Institute for Health and Care Excellence (NICE) and the full marketing authorisation for cabozantinib. The company's submission (CS) presents clinical evidence for cabozantinib and a single comparator – regorafenib; no comparison has been made against best supportive care (BSC) or any other active therapy. Cabozantinib and regorafenib are both orally administered tyrosine kinase inhibitors (TKIs); whilst these drugs are part of the same class, there are some differences in their molecular targeting profiles (further details are provided in Sections 2.4 and 2.5). The ERG's clinical advisors believe that regorafenib is the most appropriate comparator for cabozantinib. The clinical advisors also commented that the target population is small and that whilst the trial of regorafenib (RESORCE) was undertaken in a second-line population, the positive NICE recommendation for atezolizumab plus bevacizumab in the first-line setting means that regorafenib is now mostly used at third-line in people who are able to receive atezolizumab plus bevacizumab, although some patients will receive regorafenib as second-line therapy.

The CS includes a series of indirect treatment comparisons (ITCs) of cabozantinib versus regorafenib using the Bucher methodology and anchored and unanchored matching-adjusted indirect comparison (MAIC) approaches, informed by data from the pivotal trials of cabozantinib and regorafenib for HCC (CELESTIAL and RESORCE). The ITCs for progression-free survival (PFS) or overall survival (OS) indicate statistically non-significant differences in clinical outcomes between the regimens. The ITCs of safety endpoints indicate statistically non-significant differences between the regimens for individual adverse events (AEs), except for the odds of diarrhoea which was statistically significantly higher for the cabozantinib group, based on an unanchored MAIC. The CS also includes a cost-comparison analysis which suggests that, if clinical equivalence is assumed, the cost of cabozantinib (including a confidential Patient Access Scheme [PAS] discount) is less than the cost of regorafenib (excluding its comparator PAS discount).

The ERG believes that the company's case for considering cabozantinib as a Fast Track Appraisal (FTA) may not be appropriate for the following reasons:

- There is uncertainty around the treatment effect between cabozantinib and regorafenib, including the assumption of equivalence of the two regimens:

- In CELESTIAL, the OS benefit of cabozantinib over placebo was statistically significant in the second-line subgroup but not in the third-line subgroup. It was not possible to conduct ITCs in the third-line subgroup because the RESORCE trial was restricted to second-line, but regorafenib is now used in clinical practice in both second- and third-line.
- Whilst the company's ITCs consistently indicate statistically non-significant differences in PFS, OS and AEs between the regimens, the Bucher ITCs and the anchored MAICs produce point estimates of relative treatment effects which favour cabozantinib for PFS, but which favour regorafenib for OS. Both the company and the ERG prefer the anchored MAICs; however, there remain some concerns regarding the comparability of the placebo plus BSC arms of CELESTIAL and RESORCE, which means that there is uncertainty around the reliability of the results of this analysis.
- Although the ITCs for AEs indicate no statistically significant differences in individual AEs except for diarrhoea, the ERG's clinical advisors commented that cabozantinib is more toxic than regorafenib. This view is also suggested in the European Public Assessment Report (EPAR) for cabozantinib issued by the European Medicines Agency (EMA) and is likely reflected in the available Euroqol 5-Dimensions 5-Level (EQ-5D-5L) data from CELESTIAL and in the higher frequency of dose reductions in the intervention arm of CELESTIAL compared to RESORCE.

- As part of their clarification response, the company developed a partitioned survival model using PFS, OS and EQ-5D data from CELESTIAL and relative treatment effect estimates from the company's anchored and unanchored MAICs. The model was used to estimate incremental quality-adjusted life years (QALYs) for cabozantinib versus regorafenib in the second-line setting. The analyses which use relative treatment effects on PFS and OS from the anchored MAICs indicate that, excluding any toxicity-related disutilities, cabozantinib is expected to generate fewer QALYs compared with regorafenib. The company's clarification response argues that given the distribution of incremental QALY losses, there is *"no meaningful"* difference between the groups. However, the ERG notes that decisions should be made on the basis of the expectation of the mean and that the expected ICER for cabozantinib versus regorafenib would lie in the North-West or South-West quadrant, depending on the discounted prices of the products. The ERG is unsure whether the magnitude of the company's predicted incremental QALY losses are sufficient to preclude the appraisal from proceeding under the FTA route.
- The expected difference in costs for cabozantinib and regorafenib is dependent on the inclusion of PAS discounts for each product; the results of the company's cost comparison analyses including both relevant discounts cannot be reported here. These are provided in a separate confidential appendix to this report.

2. ERG'S CRITIQUE OF THE COMPANY'S DECISION PROBLEM

2.1 Introduction

The company's submission[1] (CS) presents evidence relating to the clinical effectiveness and cost of cabozantinib for adult patients with previously treated advanced unresectable hepatocellular carcinoma (HCC). The company has proposed that cabozantinib should be appraised by the National Institute for Health and Care Excellence (NICE) under its Fast Track Appraisal (FTA) process.

2.2 Health condition

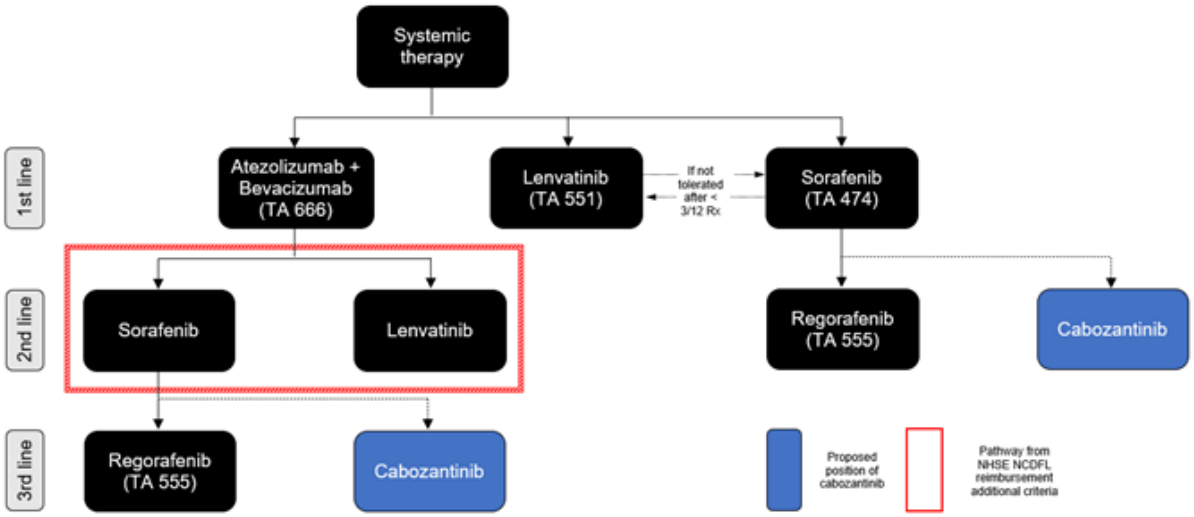
The CS[1] provides a short but accurate description of the underlying health condition. HCC is the most common form of primary liver cancer which occurs predominantly in patients with underlying chronic liver disease and cirrhosis, and is typically associated with viral hepatitis, excessive alcohol consumption, non-alcoholic steatohepatitis and haemochromatosis.[2] Based on data for the UK from 2016-2018 reported by Cancer Research UK (CRUK), there are over 6,200 new cases of liver cancer each year and around 5,600 deaths are caused by liver cancer.[3, 4] The prognosis of advanced HCC is poor with age-standardised net survival rates at 1 year and 5 years of 38.1%, and 12.7%, respectively.[4]

2.3 Current pathway for HCC and proposed positioning of cabozantinib

The company's view of the current pathway for advanced HCC and the proposed positioning of cabozantinib is shown in Figure 1. Existing NICE recommendations for treatments for advanced HCC are summarised in Table 1. The company is seeking a positive recommendation for cabozantinib in the same indication as regorafenib, which was previously appraised in NICE Technology Appraisal (TA) Number 514 (TA514) and later in TA555. In 2019, NICE recommended regorafenib as an option for treating advanced unresectable HCC in adults who have had sorafenib, only if they have Child–Pugh grade A liver impairment and an Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 or 1, and if the company provides it according to the agreed commercial arrangement.[5] The final NICE scope lists the population for the appraisal as *"Adults with advanced hepatocellular carcinoma who have had sorafenib."* If the company's target population is restricted to the same population as the NICE recommendation for regorafenib, this will be narrower than the populations defined in both the NICE scope and the marketing authorisation for cabozantinib. The company's clarification response[6] (question A1) indicates that the company would support a recommendation without restriction by Child-Pugh grade. However, the company acknowledges that only one patient in CELESTIAL[7] had Child-Pugh grade B disease and the ERG notes that the company's clinical and cost comparisons are restricted to a population in whom regorafenib is used. No comparison has been made against best supportive care (BSC) or any other active treatment (see Section 2.4).

The ERG's clinical advisors agreed that the company's view of the pathway reflects current practice. The clinical advisors commented that it is appropriate to consider cabozantinib in the same indication as that for regorafenib, as this reflects the population of patents in whom the drug would be used in practice and because it reflects the population of the CELESTIAL trial.[7] They further commented that atezolizumab plus bevacizumab has become the preferred first-line treatment for patients who are able to receive it, with sorafenib and lenvatinib now more commonly being used as second-line treatments. As regorafenib is only licensed for use after sorafenib, this treatment option is now mostly used at third-line in people who are able to receive atezolizumab plus bevacizumab, although some patients will receive regorafenib as second-line therapy. The ERG's clinical advisors commented that because atezolizumab plus bevacizumab is the preferred treatment of choice, and survival prospects in advanced HCC are poor, few patients reach third-line treatment. As such, the overall target population for cabozantinib is small. Both clinical advisors commented that they do not frequently use regorafenib.

Figure 1: Current systemic therapy treatment pathway in UK clinical practice as per NICE and Cancer Drugs Fund recommendations (reproduced from CS, Figure 2)



NCDFL - National Cancer Drug Fund List; NHSE - National Health Service England; NICE - National Institute for Health and Care Excellence; Rx – prescription

Table 1: Previous NICE recommendations for treatments for HCC

Technology	Year	Recommendation
Atezolizumab plus bevacizumab (TA666)[8]	2020	Recommended as an option for treating advanced or unresectable HCC in adults who have not had previous systemic treatment, only if: • they have Child-Pugh grade A liver impairment and an ECOG PS of 0 or 1 and • the company provides it according to the commercial arrangement.
Lenvatinib (TA551)[9]	2018	Recommended as an option for untreated, advanced, unresectable HCC in adults, only if: • they have Child–Pugh grade A liver impairment and an ECOG PS of 0 or 1 and • the company provides it according to the commercial arrangement
Sorafenib (TA474)[10]	2017	Recommended as an option for treating advanced HCC only for people with Child-Pugh grade A liver impairment, only if the company provides sorafenib within the agreed commercial access arrangement
Regorafenib (TA555)[5]	2019	Recommended as an option for treating advanced unresectable HCC in adults who have had sorafenib, only if: • they have Child–Pugh grade A liver impairment and an ECOG PS of 0 or 1 and • the company provides it according to the commercial arrangement.

TA - Technology Appraisal; HCC - hepatocellular carcinoma; ECOG - Eastern Cooperative Oncology Group; PS - performance status

2.4 Intervention

The intervention considered in the CS[1] is cabozantinib given as monotherapy. Cabozantinib is a multi-targeted inhibitor of receptor tyrosine kinases (RTKs) that potently inhibits several RTKs known to influence tumour growth, metastasis and angiogenesis, including MET, VEGFR2 and AXL.[11] Cabozantinib is available as tablets which are taken orally. The recommended daily dose of cabozantinib is 60mg per day. The marketing authorisation issued by the European Medicines Agency (EMA) is for cabozantinib as monotherapy for the treatment of HCC in adults who have previously been treated with sorafenib. The Summary of Product Characteristics (SmPC) for cabozantinib states that *"treatment should continue until the patient is no longer clinically benefiting from therapy or until unacceptable toxicity occurs."*[11] Cabozantinib is available in packs of 30 tablets at doses of 20mg, 40mg or 60mg (30 days' supply). The NHS indicative price for each pack of cabozantinib is £5,143, irrespective of the dose.[12] A Patient Access Scheme (PAS) discount is available for cabozantinib, resulting in a discounted cost per pack of [redacted] ([redacted] discount from the list price).

2.5 Comparator

The CS[1] includes a single comparator – regorafenib given as monotherapy. Regorafenib is a tumour deactivation agent that potently blocks multiple protein kinases, including kinases involved in tumour angiogenesis (VEGFR1, -2, -3, TIE2), oncogenesis (KIT, RET, RAF-1, BRAF, BRAFV600E), metastasis (VEGFR3, PDGFR, FGFR) and tumour immunity (CSF1R).[13] Regorafenib is available as tablets which are taken orally. The recommended daily dose of regorafenib is 160mg (4 x 40mg tablets)

with treatment taken for 3 weeks followed by 1 week off treatment. The EMA marketing authorisation for regorafenib is as monotherapy for the treatment of adult patients with HCC who have been previously treated with sorafenib. As with cabozantinib, the SmPC for regorafenib[13] states that treatment should continue as long as benefit is observed or until unacceptable toxicity occurs. Regorafenib is available in packs of 84 tablets at a dose of 40mg (28 days' supply). The NHS indicative price for each pack is £3,744.[12] A comparator Patient Access Scheme (cPAS) discount is available; details of this discount can be found in a separate confidential appendix to this ERG report.

The final NICE scope[14] includes a second comparator – BSC. However, BSC is not considered within the CS as it has not been recommended by NICE. The ERG agrees that BSC is not a relevant comparator for the population in whom regorafenib would otherwise be used.

2.6 Outcomes

The final NICE scope[14] lists six outcomes:

- Overall survival (OS)
- Progression-free survival (PFS)
- Response rates
- Time to treatment discontinuation (TTD)
- Adverse effects of treatment
- Health-related quality of life (HRQoL).

The pivotal study of cabozantinib for HCC is the CELESTIAL trial.[7] The pivotal study of regorafenib for HCC is the RESORCE trial.[15] The CS[1] reports data from CELESTIAL on PFS, OS, objective response rate (ORR), time on treatment and adverse events (AEs). The CS does not report data on TTD or HRQoL from CELESTIAL. The CS reports indirect treatment comparisons (ITCs) using data from the CELESTIAL and RESORCE studies[7, 15] for PFS, OS and AEs; these analyses are summarised and critiqued in Section 3 of this report. The company's cost comparison, which is underpinned by an assumption of equivalence between cabozantinib and regorafenib for all efficacy endpoints, is summarised and critiqued in Section 4 of this report.

2.7 Equality considerations

The CS[1] states that no equality issues related to the use of cabozantinib have been identified.

3. ERG'S CRITIQUE OF THE CLINICAL EFFECTIVENESS EVIDENCE SUBMITTED

3.1 Company's systematic review methods

The company conducted three searches across a wide range of sources to identify randomised controlled trials (RCTs) of cabozantinib or regorafenib in adults with advanced HCC who have received prior sorafenib (CS Appendices,[16] Section D.1.1):

1. Initial search (inception until March 2018) for cabozantinib, regorafenib, pembrolizumab, lenvatinib, nivolumab, sorafenib sunitinib, pazopanib and ramucirumab.
2. Update search (March 2018 to February 2021) for cabozantinib and regorafenib only.
3. Pragmatic search (February 2021 until January 2022) by applying high specificity RCT filters.

Two RCTs met the inclusion criteria: the CELESTIAL trial of cabozantinib[7] and the RESORCE trial of regorafenib.[15] Despite the differences between the three searches, the ERG is not aware of any relevant RCTs for cabozantinib and regorafenib that have been missed. Both trials are summarised side-by-side in the following sections.

3.2 Summary and critique of CELESTIAL and RESORCE trials

3.2.1 Overview of trials

The CS[1] focusses on a comparison between two trials: the CELESTIAL trial of cabozantinib plus BSC versus placebo plus BSC,[7] and the RESORCE trial of regorafenib plus BSC versus placebo plus BSC.[15] The two trials are summarised in Table 2, based on data presented in the CS on CELESTIAL (CS, Section B.3), RESORCE (CS Appendices,[16] Section D.1.1.9) and both trials (CS, Table 22). Patients in both trials had received prior sorafenib. CELESTIAL included both second- and third-line patients, whereas RESORCE included only second-line patients. RESORCE only included patients who had tolerated sorafenib, whereas CELESTIAL included patients irrespective of tolerance of sorafenib.

3.2.2 Study quality of CELESTIAL and RESORCE trials

The CS[1] presents a quality assessment of CELESTIAL[7] and RESORCE[15] using the standard NICE criteria for RCTs (CS, Section B.3.5 and CS Appendices, Section D.1.3). Both trials were considered to be of good methodological quality on all criteria. The ERG agrees with this assessment.

Table 2: Summary of design of CELESTIAL and RESORCE trials

Trial name	CELESTIAL	RESORCE
Intervention (N)	Cabozantinib (60mg/day) plus BSC (N=470)	Regorafenib (160mg/day, weeks 1-3 per 4-week cycle) plus BSC (N=379)
Comparator (N)	Placebo plus BSC (N=237)	Placebo plus BSC (N=194)
Analysis sets: **- ITT (all randomised)** **- Safety (≥1 dose)**	ITT: Cabozantinib 470, Placebo 237 Safety: Cabozantinib 467, Placebo 237	ITT: Regorafenib 379, Placebo 194 Safety: Regorafenib 374, Placebo 193
Patient population; key inclusion criteria (ITT)	- Second and third-line patients - Received prior sorafenib - Sorafenib tolerant and intolerant - Progression following ≥1 prior systemic treatment - ECOG PS 0 or 1 - Child-Pugh status A (further inclusion criteria: CS, Table 9)	- Second-line patients only - Failure on prior sorafenib - Sorafenib tolerant only - ECOG PS 0 or 1 - Child-Pugh status A (further inclusion criteria: CS Appendix D.1.1.9, Table 12)
Methodology	Phase III, double-blind	Phase III, double-blind
Stratification factors	- Aetiology of disease (hepatitis B, hepatitis C, other) - Geographic region (Asia, other) - Extrahepatic disease and/or macrovascular invasion (yes, no)	- Geographical region (Asia, other) - Extrahepatic disease (yes, no) - Macrovascular invasion (yes, no) - α-fetoprotein (<400 or ≥400 ng/mL) - ECOG PS (0, 1)
Study initiation and completion (years)	September 2013 – June 2017 (data cut-off date)	May 2013 – Feb 2016 (primary completion date)
Study centres	- Multicentre (Europe, North America, Australia, New Zealand, Asia) - UK: [redacted]	- Multicentre (Europe, North America, Australia, South America, Asia) - UK: 5 study sites, 20 participants
Treatment stopping rule	Continued as long as patient had clinical benefit (as judged by investigator) or until unacceptable toxicity[7, 17]	Continued until disease progression as defined by mRECIST, clinical progression (defined as an ECOG PS ≥3 or symptomatic deterioration, including increased liver function tests), death, unacceptable toxicity, withdrawal of consent by the patient, or decision by the treating physician that discontinuation would be in the patient's best interest.[15]
Median follow-up	22.9 months (2017 data-cut)	7.0 months (2016 data-cut)
% censored for OS	32%	37%
Outcomes	- OS - PFS: via RECIST 1.1 - TTD - ORR: complete or partial response - HRQoL: EQ-5D-5L until 8 weeks after progression or discontinuation - Safety and tolerability	- OS - PFS: via RECIST 1.1 and mRECIST - TTP - ORR: complete or partial response - HRQoL: EQ-5D - Safety and tolerability

BSC - best supportive care; ECOG - Eastern Cooperative Oncology Group; PS - performance status; HRQoL - health-related quality of life; ITT - intention-to-treat; N - number of participants; ORR - overall response rate; OS - overall survival; PFS - progression-free survival; RECIST - Response Evaluation Criteria in Solid Tumours; TTD - time to treatment discontinuation; TTP - time to progression

3.2.3 Baseline characteristics: CELESTIAL and RESORCE

The baseline characteristics of CELESTIAL[7] and RESORCE[15] are shown in Table 3. This table is based on data presented in CS,[1] Table 10 (CELESTIAL), CS Appendix D,[16] Table 13 (RESORCE) and CS, Table 23 (both trials).

Comparison of baseline characteristics between trials: The CELESTIAL and RESORCE trials[7, 15] were similar in terms of age, sex, Child-Pugh grade, baseline disease spread, aetiology (hepatitis B and C or alcohol-related) and alpha-fetoprotein. Key differences were as follows. CELESTIAL patients were 72% second-line and 28% third-line, whereas RESORCE patients were entirely second-line. CELESTIAL included patients irrespective of whether they had tolerated sorafenib, whereas RESORCE included only patients who had tolerated sorafenib. The European Public Assessment Report (EPAR) for cabozantinib[11] states that 96% of patients in CELESTIAL had progressed on prior sorafenib and *"therefore, it seems unlikely that many sorafenib-intolerant patients were recruited"*. Patients in CELESTIAL had a shorter duration of prior sorafenib treatment (mean of 8 versus 12 months). CELESTIAL had fewer Asian patients than RESORCE (34% versus 41%), more white patients (56% versus 36%), and fewer patients from the Asian geographic region (25% versus 38%). ECOG PS was slightly worse in CELESTIAL (53% ECOG PS 0, 47% ECOG PS 1) than RESORCE (66% ECOG PS 0, 34% ECOG PS 1). In terms of prognosis, the EPAR for cabozantinib[11] states that *"there are no important differences between the two trial populations that may have impacted efficacy."* The ERG's clinical advisors stated that patients in RESORCE may have had a better prognosis as they were all second-line. Conversely however, the ERG's advisors also suggested that line of treatment may make little difference since other prognostic factors were similar between the trials, and one advisor further commented that patients reaching third-line treatment would have a better disease biology than those at second-line, by virtue of reaching this line of therapy. The clinical advisors considered that the restriction to sorafenib-tolerant patients in RESORCE was unlikely to substantially affect prognosis.

Relevance of trials to UK HCC population: The ERG's clinical advisors stated that the CELESTIAL trial population did not reflect the full UK population of advanced HCC post-sorafenib patients as it restricted the population to those with ECOG PS 0-1 and Child-Pugh grade A. However, the clinical advisors considered that the trial reflected the population of patients who are likely to receive cabozantinib in clinical practice, as patients would need to be relatively fit in order to tolerate it. The CS[1] reports a comparison of the CELESTIAL trial population versus a retrospective UK audit of 448 advanced HCC patients from 15 hospitals having received first-line sorafenib[18] (CS, Table 11). Patient characteristics were broadly similar, though more patients in CELESTIAL (versus those in the UK audit) had ECOG PS 0 and Child-Pugh grade A, and more patients in CELESTIAL had extrahepatic spread or hepatitis B or C.

Table 3: Baseline characteristics in CELESTIAL and RESORCE

	CELESTIAL		RESORCE	
Treatment (N)	**Cabozantinib (N = 470)**	**Placebo (N=237)**	**Regorafenib (N = 379)**	**Placebo (N=194)**
Age, years: median (range)	64 (22-86)	64 (24-86)	64 (54-71)	62 (55-68)
Male (%)	81	85	88	88
Race (%)				
White	56	55	36	35
Asian	34	35	41	40
Other	10	10	23	25
Geographic region				
Europe	49	46	NR	NR
Asia	25	25	38	38
USA/Canada	23	25	NR	NR
Australia/New Zealand	3	5	NR	NR
ECOG PS (%)				
0	52	55	65	67
1	48	45	35	33
Child-Pugh status (%)				
A	98	99	98	97
B	1	0.8	1	3
Baseline disease (%)				
Extrahepatic spread	79	77	70	76
Macrovascular invasion	27	34	29	28
Aetiology at baseline (%)				
Hepatitis B	38	38	38	38
Hepatitis C	24	23	21	21
Alcohol-related	24	16	24	28
NASH	9	10	7	7
Other/unknown	21	27	24	21
Alpha-fetoprotein ≥400 ng/mL (%)	41	43	43	45
Line of treatment (systemic):				
Second	71	73	100	100
Third	28	26	0	0
Duration prior sorafenib, months				
Mean	8	NR	12	NR
Median	5.3	4.8	NR	NR
Range	0.3 to 70.0	0.2 to 76.8	NR	NR
Time from progression on sorafenib (as most recent systemic agent), months, median	1.61	1.66	NR	NR
Prior local therapy (inc. TACE) (%)	44	48	NR	NR
Prior TACE (%)	43	47	NR	NR

AFP - alpha-fetoprotein; ECOG - Eastern Cooperative Oncology Group; ITT - intention to treat; NASH - non-alcoholic steatohepatitis; TACE - transarterial chemoembolisation

3.2.4 Clinical effectiveness: OS and PFS (CELESTIAL and RESORCE)

Results for OS and PFS for CELESTIAL[7] and RESORCE[15] are summarised in Table 4, which presents medians, hazard ratios (HRs) and 95% confidence intervals (CIs) for intention-to-treat (ITT) analyses

and subgroups by line of therapy. These results are based on data presented in CS,[1] Section B.3.6 (CELESTIAL), CS Appendix E[16] (CELESTIAL subgroups) and CS Appendix D.1.1.9 (RESORCE).

OS: The Kaplan-Meier plot for OS in CELESTIAL[7] is shown in Figure 2. In CELESTIAL, there was a statistically significant difference in OS between cabozantinib and placebo in the ITT population at the 2017 data cut-off (HR 0.76, 95% CI 0.63 to 0.92) and in the second-line subgroup (HR 0.74, 95% CI 0.59 to 0.92), but not in the third-line subgroup (HR 0.90, 95% CI 0.63 to 1.29). The company's clarification response[19] (question A10) highlights the lower patient numbers in the third-line subgroup (28% of trial patients) and notes that regorafenib is currently being used as third-line treatment in NHS practice, despite the lack of trial evidence. In RESORCE,[15] there was a statistically significant difference in OS between regorafenib and placebo in the second-line ITT population, both at the 2016 data cut-off (HR 0.63, 95% CI 0.50 to 0.79) and at later cut-offs (see Table 4).

The CS[1] notes that some patients in CELESTIAL[7] and RESORCE[15] continued to receive their assigned treatment beyond disease progression. The company's clarification response[19] (questions A11 and B2) states that this was more pronounced for regorafenib and that this may bias OS in favour of regorafenib. The clarification response (question B2) also states that subsequent systemic anticancer therapies were received by 25% of the cabozantinib arm in CELESTIAL and 23.2% of the regorafenib arm in RESORCE. The company states that since the numbers were relatively small and similar across trials, the effect of subsequent treatment on OS is expected to be limited.

PFS: The Kaplan-Meier plot for PFS in CELESTIAL[7] is shown in Figure 3. In CELESTIAL, there was a statistically significant difference in PFS between cabozantinib and placebo in the ITT population at the 2017 data cut-off (HR 0.44, 95% CI 0.36 to 0.52) and in the second-line subgroup (HR 0.43, 95% CI 0.35 to 0.52), while in the third-line subgroup, results were less favourable though still statistically significant (HR 0.58, 95% CI 0.41 to 0.83). In RESORCE,[15] PFS was statistically significant in the second-line ITT population at the 2016 data cut-off, both when using RECIST 1.1 (HR 0.43, 95% CI 0.35 to 0.52) and modified RECIST (mRECIST) (HR 0.46, 95% CI 0.37 to 0.56).

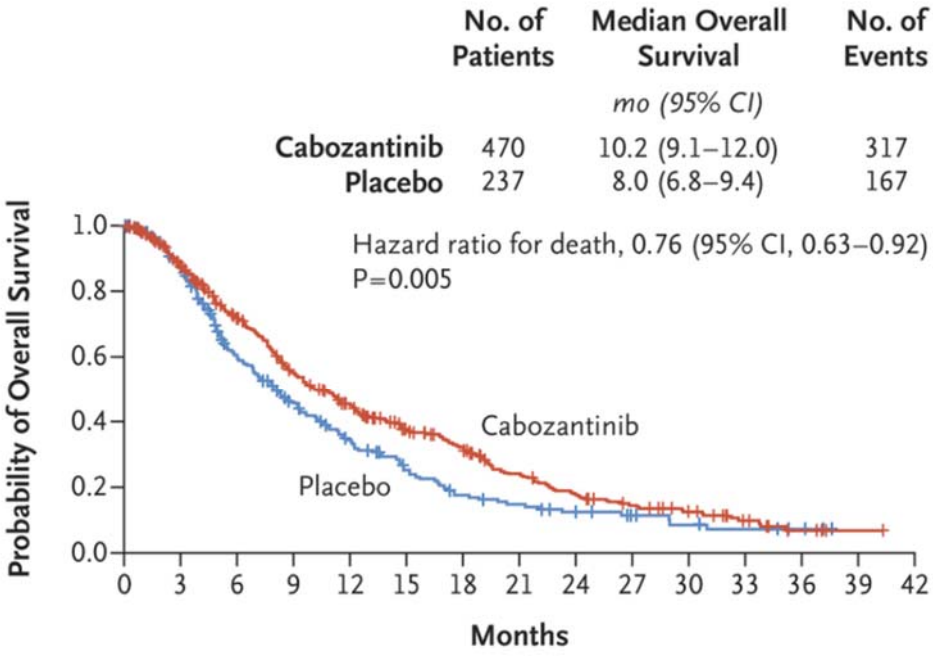
OS and PFS data used in ITC: Table 4 also indicates which data were used in the company's indirect treatment comparisons (ITCs), which include Bucher ITCs and matching-adjusted indirect comparisons (MAICs). The company ITCs are detailed further in Section 3.3 of this report.

Table 4: OS and PFS: CELESTIAL and RESORCE

Line of treatment	Criteria	CELESTIAL					RESORCE				
		Data-cut (FU)	Cabozantinib: median	Placebo: median	HR (95% CI)	Used in analysis	Data-cut (FU)	Regorafenib: median	Placebo: median	HR (95% CI)	Used in analysis
OS											
Second 72% Third 28%		2017 (22.9mo) **(ITT)**	10.2 months	8.0 months	0.76 (0.63 to 0.92)	Bucher ITT					
Second		2017 (22.6mo)	11.4 months	7.7 months	0.74 (0.59 to 0.92)	Bucher 2L MAIC	2016 (7.0mo)	10.6 months	7.8 months	0.63 (0.50 to 0.79)	Bucher ITT MAIC
							2017 (NR)	10.7 months	7.9 months	0.61 (0.50 to 0.75)	
							2018 (NR)	10.7 months	7.9 months	0.62 (0.51 to 0.75)	Bucher 2L
Third		2017 (NR)	8.6 months	8.6 months	0.90 (0.63 to 1.29)						
PFS											
Second 72% Third 28%	RECIST 1.1	2017 (22.9mo) **(ITT)**	5.2 months	1.9 months	0.44 (0.36 to 0.52)	Bucher ITT					
Second	RECIST 1.1	2017 (22.6mo)	5.5 months	1.9 months	0.43 (0.35 to 0.52)	Bucher 2L MAIC	2016 (7.0mo)	3.4 months	1.5 months	0.43 (0.35 to 0.52)	Bucher 2L
	mRECIST						2016 (7.0mo)	3.1 months	1.5 months	0.46 (0.37 to 0.56)	Bucher ITT MAIC
Third	RECIST 1.1	2017 (NR)	3.7 months	1.9 months	0.58 (0.41 to 0.83)						

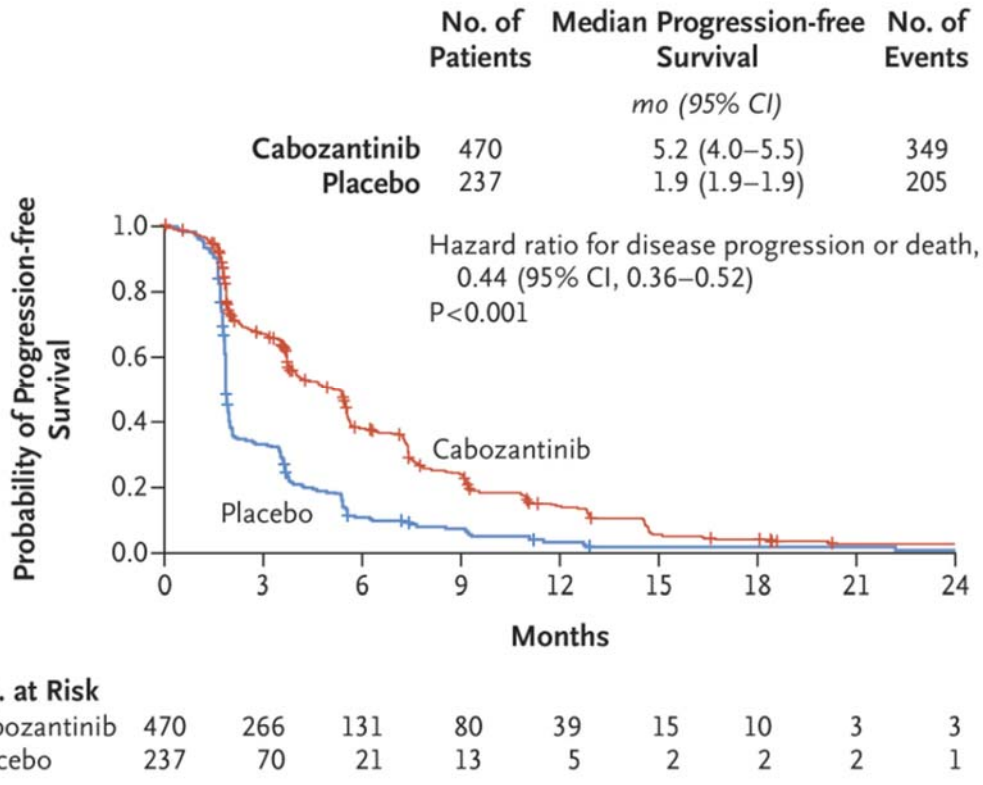
Bucher ITT = CELESTIAL 2nd/3rd-line vs. RESORCE 2nd-line (presented in CS); Bucher 2L = all 2nd-line (presented in company's clarification response,[19] question A13).
CI - confidence interval; FU- follow-up; HR - hazard ratio; ITT - intention-to-treat; MAIC - matching-adjusted indirect comparison; mo - months; NR - not reported; OS - overall survival; PFS - progression-free survival; RECIST - Response Evaluation Criteria in Solid Tumours
Source: CS,[1] Tables 16 and 17 (CELESTIAL), CS Appendix E,[16] Table 22 (CELESTIAL subgroups) and CS Appendix D, Table 14 (RESORCE)

Figure 2: Kaplan-Meier plot for OS, CELESTIAL (ITT, 2017 data cut-off)



Source: Abou-Alfa et al., 2018 [17].

Figure 3: Kaplan-Meier plot for PFS, CELESTIAL (ITT, 2017 data cut-off)



Source: Abou-Alfa et al., 2018 [17].

3.2.5 Subgroup analyses for OS and PFS (CELESTIAL and RESORCE)

The subgroup analyses for OS and PFS in CELESTIAL[7] and RESORCE[15] are provided in CS Appendix E, Figure 6 and CS Appendix D, Figure 3.[16] In CELESTIAL, the HRs for PFS and OS for cabozantinib were less favourable for third-line than second-line patients. In addition, the HR for OS in CELESTIAL was close to 1.0 (indicating little effect of cabozantinib) in patients from Asia, patients without extrahepatic disease, and patients with hepatitis C virus. The company's clarification response[19] (question A9) states that clinical experts who attended an advisory board held by the company considered these findings to be related to small sample sizes as there was no clear clinical explanation. In RESORCE, the HR for OS was close to 1.0 (indicating little effect of regorafenib) in patients without extrahepatic disease and patients with a history of alcohol use.

3.2.6 Clinical effectiveness: Overall response rate (CELESTIAL and RESORCE)

The overall response rates (ORRs) for CELESTIAL[7] and RESORCE[15] are shown in Table 5 (data from CS,[1] Table 19 and CS Appendix,[16] Table 14). Using RECIST 1.1, the ORR in CELESTIAL was 4% for cabozantinib and 0.4% for placebo, whilst the ORR in RESORCE was 7% for regorafenib and 3% for placebo. All were partial responses (PR); there were no complete responses (CR) in either trial when using RECIST 1,1.

Table 5: Overall response rate in CELESTIAL and RESORCE

Response: n (%)	**CELESTIAL (RECIST 1.1)**		**RESORCE (RECIST 1.1)**		**RESORCE (mRECIST)**	
	Cabozantinib (N = 470)	**Placebo (N=237)**	**Regorafenib (N = 379)**	**Placebo (N=194)**	**Regorafenib (N = 379)**	**Placebo (N=194)**
ORR [CR+PR]	18 (4%)	1 (0.4%)	25 (7%)	5 (3%)	40 (11%)	8 (4%)
CR	0	0	0	0	2 (0.5%)	0
PR	18 (4%)	1 (0.4%)	25 (7%)	5 (3%)	38 (10%)	8 (4%)

Source: CELESTIAL: CS,[1] Table 19; RESORCE: CS Appendix,[16] Table 14 and Bruix et al., 2017[15]
CR - complete response; ORR - overall response rate; PR - partial response; RECIST - Response Evaluation Criteria in Solid Tumours

3.2.7 HRQoL (CELESTIAL and RESORCE)

HRQoL in CELESTIAL: HRQoL data for CELESTIAL[7] are not presented in the CS or its appendices.[1, 16]

[redacted]

[redacted]

[redacted] The ERG's clinical advisors stated that HRQoL is a very important factor in this population and that there is a need to consider the balance between positive gains of treatment in PFS and OS and negative effects on HRQoL.

Table 6: CELESTIAL: EQ-VAS and EQ-Index Scores: Change from baseline, repeated-measures analysis (EQ-5D Index: ITT population for countries in which index is validated; EQ-VAS: ITT population)

	Cabozantinib (N = 470) LS means (SE) [n]	Placebo (N = 237) LS means (SE) [n]	Difference in mean change[a]	Pooled SD	*p*-value[a]	Effect size[b]
EQ-5D index	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-VAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

HRQoL in RESORCE: The NICE TA555 guidance for regorafenib[5] states that HRQoL scores were generally similar across treatment arms with different measures, including the EQ-5D. Scores were slightly worse for regorafenib than for BSC but these differences did not pass the 'minimally important difference' threshold established in the literature. The TA555 guidance also states that the EQ-5D utility values from RESORCE[15] appear high for patients who have progressed on sorafenib, and that most patients tend to have side effects from treatment with a serious impact on their HRQoL, which did not appear to be reflected in the utility values. The EQ-5D decrement for progression (–0.048) in RESORCE appeared low for an advanced HCC population with progressed disease. It was also noted that the EQ-5D questionnaire was completed on the first day of each treatment cycle, when a patient had not had treatment for a week.

3.2.8 Safety (CELESTIAL and RESORCE)

Adverse event (AE) data are provided for CELESTIAL[7] in the CS,[1] Section B.3.8 (Tables 20 and 21) and CS Appendix F,[16] and for RESORCE[15] in the CS Appendix D.1.1.9 (Table 16). During the clarification stage, the company provided summary data on AEs for both CELESTIAL and RESORCE. These data are provided in Table 7 and Table 8.

Safety overview for CELESTIAL: In CELESTIAL,[7] AEs occurred as follows for cabozantinib vs. placebo (Table 7): Grade 3 or 4 AEs (68% vs. 36%); serious adverse events (SAEs) (50% vs. 37%); treatment-related SAEs (18% vs. 6%); AEs leading to dose modification (89% vs. 40%) and AEs leading to discontinuation (21% vs. 4%).

Comparison with RESORCE: An overview of AEs for RESORCE[15] is also shown in Table 7. The percentages of Grade 3 or 4 AEs appeared similar in CELESTIAL and RESORCE, while SAEs and treatment-related SAES appeared slightly higher in CELESTIAL than in RESORCE. AEs leading to dose modification also appeared somewhat higher in CELESTIAL, while AEs leading to discontinuation appeared similar in the two active treatment arms, though the difference from placebo was more marked in CELESTIAL.

Table 7: Summary of safety data in CELESTIAL and RESORCE

AEs	Cabozantinib (n=467), n (%)	Placebo (n=237), n (%)	Regorafenib (N=374), n (%)	Placebo (N=193), n (%)
Any AE (all grades)	460 (99)	219 (92)	374 (100)	179 (93)
Treatment-related AEs	439 (94)	148 (62)	346 (93)	100 (52)
Grade 3 or 4 AEs	316 (68)	86 (36)	248 (66)	75 (38)
SAEs	232 (50)	87 (37)	166 (44)	90 (47)
Treatment-related SAEs	82 (18)	14 (5.9)	36 (10)	5 (3)
Treatment-related Grade 5 AEs (deaths)	6 (1.3)	1 (0.4)	7(2)	2 (1)
Deaths (at any time, excluding PD)	314 (67)	167 (70)	50 (13)	38 (20)
AE leading to dose modification	416 (89)	94 (40)	255 (68)	60 (31)
AE leading to discontinuation of study drug	96 (21)	10 (4.2)	93 (25)	37 (19)

AEs - adverse events; PD - progressive disease; SAEs - serious adverse events
Source: CS,[1] Table 20 (CELESTIAL), CS Appendices,[16] Section D.1.1.9 (RESORCE) and company's clarification response[6] (question A7)

Individual AEs for CELESTIAL: In CELESTIAL,[7] the most common AEs (see Table 8) were as follows (for cabozantinib vs. placebo): diarrhoea (54% vs. 44%); decreased appetite (48% vs. 18%); palmar-plantar erythrodysaesthesia syndrome (PPES or hand-foot syndrome) (46% vs. 5%); fatigue (45% vs. 30%); nausea (31% vs. 18%); hypertension (29% vs. 6%); vomiting (26% vs. 12%); increased aspartate aminotransferase (AST) (22% vs. 11%) and asthenia (22% vs. 8%). The most common Grade 3 or 4 AEs were: PPES (17%, vs. 0%); hypertension (16% vs. 2%); increased AST (12% vs. 7%); fatigue (10% vs. 4%) and diarrhoea (10% vs. 2%). Treatment-related deaths occurred in 6 patients in the cabozantinib arm (hepatic failure, tracheoesophageal fistula, portal-vein thrombosis, upper gastrointestinal haemorrhage, pulmonary embolism, hepatorenal syndrome) and in 1 patient in the

placebo arm (hepatic failure). The CS[1] states that AEs with cabozantinib were typical of those with TKI therapies.

Comparison with RESORCE: AE data from RESORCE[15] for regorafenib versus placebo are also presented in Table 8. Section B.3.10 of the CS[1] presents the results of ITCs between cabozantinib and regorafenib for selected AEs. The company ITCs are discussed further in Section 3.3 of this report.

The EPAR for cabozantinib[11] (page 106) states that, based on reported safety data for both drugs, *"cabozantinib appears to be more toxic than regorafenib."* The ERG's clinical advisors were asked about their views on comparative toxicity of cabozantinib and regorafenib. One advisor stated that, based on their clinical experience and the trial results, they considered cabozantinib to have a more severe and less predictable AE profile than regorafenib, with many patients on cabozantinib requiring dose reductions or discontinuation due to AEs (key AEs impacting on patients, based on their experience, included diarrhoea, severe fatigue and mouth ulcers). The other clinical advisor did not have experience of using cabozantinib, but noted that trial results suggested higher levels of AE-related dose modifications with cabozantinib than regorafenib. One of the clinical advisors commented that the inclusion of sorafenib-intolerant patients may have contributed to the higher numbers of AEs in CELESTIAL[7] than RESORCE.[15] However, as noted in Section 3.2.3, the EPAR for cabozantinib[11] states that 96% of patients in CELESTIAL had progressed on previous sorafenib and *"therefore, it seems unlikely that many sorafenib-intolerant patients were recruited."* The ERG's clinical advisor with experience of using the drug considered that the higher number of AEs for cabozantinib was likely to be attributable to its different mechanism of action.

Table 8: AEs (any grade) reported in ≥10% of patients in either treatment group for CELESTIAL and RESORCE

AEs	Cabozantinib (n=467) n (%)			Placebo (n=237) n (%)			Regorafenib (N=374) n (%)			Placebo (N=193) n (%)		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
Any AE	460 (99)	270 (58)	46 (10)	219 (92)	80 (34)	6 (3)	374 (100)	208 (56)	40 (11)	179 (93)	61 (32)	14 (7)
Diarrhoea	251 (54)	45 (10)	1 (<1)	44 (19)	4 (2)	0	155 (41)	12 (3)	0	29 (15)	0	0
Decreased appetite	225 (48)	27 (6)	0	43 (18)	1 (<1)	0	NR	NR	NR	NR	NR	NR
PPES	217 (46)	79 (17)	0	12 (5)	0	0	198 (53)	47 (13)	NA	15 (8)	1 (1)	NA
Fatigue	212 (45)	49 (10)	0	70 (30)	10 (4)	0	151 (40)	34 (9)	NA	61 (32)	9 (5)	NA
Nausea	147 (31)	10 (2)	0	42 (18)	4 (2)	0	64 (17)	2 (1)	NA	26 (13)	0	NA
Hypertension	137 (29)	73 (16)	1 (<1)	14 (6)	4 (2)	0	116 (31)	56 (15)	1 (<1)	12 (6)	9 (5)	0
Vomiting	121 (26)	2 (<1)	0	28 (12)	6 (3)	0	47 (13)	3 (1)	0	13 (7)	1 (1)	0
Increase in AST level	105 (22)	51 (11)	4 (1)	27 (11)	15 (6)	1 (<1)	92 (25)	37 (10)	4 (1)	38 (20)	19 (10)	3 (2)
Asthenia	102 (22)	31 (7)	1 (<1)	18 (8)	4 (2)	0	NR	NR	NR	NR	NR	NR
Dysphonia	90 (19)	3 (1)	0	5 (2)	0	0	NR	NR	NR	NR	NR	NR
Constipation	87 (19)	2 (<1)	0	45 (19)	0	0	65 (17)	1 (<1)	0	22 (11)	1 (1)	0
Abdominal pain	83 (18)	7 (1)	1 (<1)	60 (25)	10 (4)	0	105 (28)	13 (3)	NA	43 (22)	8 (4)	NA
Weight loss	81 (17)	5 (1)	0	14 (6)	0	0	51 (14)	7 (2)	NA	9 (5)	0	NA
Increase in ALT level	80 (17)	23 (5)	0	13 (5)	5 (2)	0	55 (15)	10 (3)	2 (1)	22 (11)	5 (3)	0
Mucosal inflammation†	65 (14)	8 (2)	0	5 (2)	1 (<1)	0	47 (13)	4 (1)	0	6 (3)	1 (1)	0
Pyrexia	64 (14)	0	0	24 (10)	1 (<1)	0	72 (9)	0	0	14 (7)	0	0
Upper abdominal pain	63 (13)	3 (1)	0	31 (13)	0	0	NR	NR	NR	NR	NR	NR
Cough	63 (13)	1 (<1)	0	26 (11)	0	0	40 (11)	1 (<1)	NA	14 (7)	0	NA
Peripheral oedema**	63 (13)	4 (1)	0	32 (14)	2 (1)	0	60 (16)	2 (1)	NA	24 (12)	0	NA
Stomatitis	63 (13)	8 (2)	0	5 (2)	0	0	NR	NR	NR	NR	NR	NR
Dyspnoea	58 (12)	15 (3)	0	24 (10)	1 (<1)	0	NR	NR	NR	NR	NR	NR
Rash	58 (12)	2 (<1)	0	14 (6)	1 (<1)	0	NR	NR	NR	NR	NR	NR
Ascites	57 (12)	17 (4)	1 (<1)	30 (13)	11 (5)	0	58 (16)	16 (4)	0	31 (16)	11 (6)	0
Dysgeusia	56 (12)	0	0	5 (2)	0	0	NR	NR	NR	NR	NR	NR
Hypoalbuminemia	55 (12)	2 (<1)	0	12 (5)	0	0	57 (15)	6 (2)	0	16 (8)	1 (1)	0
Headache	52 (11)	1 (<1)	0	16 (7)	1 (<1)	0	NR	NR	NR	NR	NR	NR
Insomnia	49 (10)	1 (<1)	0	17 (7)	0	0	NR	NR	NR	NR	NR	NR

AEs	Cabozantinib (n=467) n (%)			Placebo (n=237) n (%)			Regorafenib (N=374) n (%)			Placebo (N=193) n (%)		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
Dizziness	48 (10)	2 (<1)	0	15 (6)	0	0	NR	NR	NR	NR	NR	NR
Dyspepsia	47 (10)	0	0	7 (3)	0	0	NR	NR	NR	NR	NR	NR
Anaemia	46 (10)	18 (4)	1 (<1)	19 (8)	12 (5)	0	58 (16)	16 (4)	2 (1)	22 (11)	0	NA
Back pain	46 (10)	5 (1)	0	24 (10)	1 (<1)	0	42 (11)	6 (2)	1 (<1)	17 (9)	2 (1)	0
Increase in serum bilirubin level	45 (10)	10 (2)	4 (1)	17 (7)	2 (1)	2 (1)	108 (29)	37 (10)	2 (1)	34 (19)	15 (8)	6 (3)
Decrease in platelet count	45 (10)	13 (3)	4 (1)	7 (3)	2 (1)	0	29 (10)	13 (3)	1 (<1)	5 (3)	0	0

† Mucosal inflammation reported in CELESTIAL, whereas in RESORCE oral mucositis reported
*** Peripheral oedema reported in CELESTIAL, whereas in RESORCE limb oedema recorded.*
AE - adverse event; ALT - alanine aminotransferase; AST - aspartate aminotransferase; NA - not applicable; NR - not reported; PPES - palmar-plantar erythrodysaesthesia syndrome.
Source: CS,[1] Table 21 (CELESTIAL), CS Appendices,[16] Section D.1.1.9 Table 16 (RESORCE) and company's clarification response[19] (question A7, Table 3)

3.2.9 Ongoing studies of cabozantinib and regorafenib

The CS[1] states that no relevant studies of cabozantinib for advanced HCC are expected to report in the next 12 months. The company's clarification response[19] (question A8) states that there are no ongoing or planned studies of cabozantinib or regorafenib in the post-sorafenib setting.

3.2.10 Summary of ERG's critique of clinical effectiveness evidence

The main points highlighted by the ERG relating to the clinical effectiveness evidence are as follows:

- Population: The CS[1] focusses on patients with advanced HCC who have had prior sorafenib and have Child–Pugh grade A liver impairment and an ECOG PS of 0 or 1. The ERG's clinical advisors considered this appropriate as it is consistent with the populations of the relevant trials[7, 15] and reflects the population who would be treated in clinical practice.
- Clinical trials: The CS focusses on a comparison between the CELESTIAL trial of cabozantinib and the RESORCE trial of regorafenib. Patients in both trials had received prior sorafenib. Almost all trial patients had Child-Pugh grade A and ECOG PS 0-1. CELESTIAL included both second- and third-line patients while RESORCE included only second-line patients. RESORCE included sorafenib-tolerant patients only, while CELESTIAL included patients irrespective of sorafenib tolerance. The ERG's clinical advisors were uncertain to what extent these differences would affect PFS, OS and AEs.
- OS: CELESTIAL showed a statistically significant effect of cabozantinib on OS in the ITT population and in the second-line subgroup, but not in the third-line subgroup. In RESORCE, there was a statistically significant effect of regorafenib on OS in the second-line ITT population, whilst there is no RCT evidence in third-line patients.
- PFS: CELESTIAL showed a statistically significant effect of cabozantinib on PFS in the ITT population and in the second-line and third-line subgroups, though results were less favourable in the third-line subgroup. In RESORCE, the treatment effect on PFS was statistically significant in the second-line ITT population.
- HRQoL: The mean difference in change from baseline EQ-5D for regorafenib versus placebo in RESORCE was reported to be small and non-significant. [redacted].
- Safety: In CELESTIAL, AEs occurred as follows for cabozantinib vs. placebo: Grade 3 or 4 AEs (68% vs. 36%); SAEs (50% vs. 37%), treatment-related SAEs (18% vs. 6%); AEs leading to dose modification (89% vs. 40%) and AEs leading to discontinuation (21% vs. 4%). The most common AEs were: diarrhoea; decreased appetite; PPES; fatigue; nausea; hypertension; vomiting; increased AST and asthenia. The ERG's clinical advisors considered cabozantinib to have a more severe AE profile than regorafenib. One of the ERG's clinical advisors believed

that it is not clear to what extent the trial AE results were affected by the inclusion of sorafenib-intolerant patients in CELESTIAL. One of the ERG's clinical advisors commented that the reason for not including sorafenib-intolerant patients in RESORCE was because regorafenib is essentially the same molecule as sorafenib, but that cabozantinib is different. As noted in Section 3.2.2, the EPAR for cabozantinib suggests that it is unlikely that many sorafenib-intolerant patients were recruited into CELESTIAL.

3.3 Summary and critique of company's indirect treatment comparisons

3.3.1 Summary of ITCs presented

As discussed in Section 2.3, the company is seeking a positive NICE recommendation for cabozantinib which is the same as that for regorafenib. Owing to the absence of direct evidence comparing cabozantinib against regorafenib, the CS[1] (Section B.3.10) presents the results of a series of ITCs of these treatments. These ITCs utilise data from the CELESTIAL and RESORCE trials.[7, 15] The ERG agrees that ITC methods are required to provide estimates of relative treatment effects.

ITC analyses are presented in the CS[1] and the company's clarification response[19] for OS, PFS and a number of individual safety endpoints (Grade 3/4 AEs which occurred in ≥5% of patients in either arm), including: increased AST; elevated bilirubin; fatigue; hypertension; diarrhoea and PPES. The ITC analyses submitted by the company comprise:

- ITCs using the Bucher approach,[20] comparing cabozantinib against regorafenib, anchoring on placebo plus BSC (which is used as the common comparator arm) using aggregate level data from both the CELESTIAL and RESORCE trials.[7, 15]
- Anchored MAICs comparing cabozantinib against regorafenib (using placebo plus BSC as a common comparator arm), using individual patient data (IPD) from the CELESTIAL trial. This analysis relies upon the assumption of proportional hazards (PH), and uses a Cox PH model to estimate a constant HR.
- Anchored MAICs comparing cabozantinib and regorafenib (using placebo plus BSC as a common comparator arm), using IPD from the CELESTIAL trial. This analysis does not rely upon the PH assumption, and instead involved fitting a series of independent parametric models to both arms of the weighted CELESTIAL and RESORCE trials to estimate a time-varying HR.
- Unanchored MAICs comparing cabozantinib against regorafenib by comparing absolute treatment effects by fitting independent parametric models to the weighted cabozantinib arm from CELESTIAL and the regorafenib arm from RESORCE.

A summary of the ITC analyses conducted by the company is presented in Table 9.

Table 9: Summary of company's ITC analyses

ITC method	Study	Population	Arms utilised in comparison	Type of comparison	Attempts to adjust for imbalances in trial populations	Outcomes assessed
Bucher indirect comparison	CELESTIAL	ITT: second- and third-line HCC patients	Cabozantinib; placebo plus BSC	Anchored	No	Efficacy: OS; PFS Safety: Hypertension; elevated AST; fatigue
	RESORCE	ITT: second-line HCC patients	Regorafenib; placebo plus BSC			
Bucher indirect comparison*	CELESTIAL	Subpopulation: second-line HCC patients	Cabozantinib; placebo plus BSC	Anchored	No	Efficacy: OS; PFS
	RESORCE	ITT: second-line HCC patients	Regorafenib; placebo plus BSC			
MAIC using constant HR (Cox PH model for OS/PFS) or OR (safety outcomes)	CELESTIAL	Subpopulation: second-line HCC patients	Cabozantinib; placebo plus BSC	Anchored	Yes	Efficacy: OS; PFS Safety: Increased AST, elevated bilirubin; fatigue; hypertension
	RESORCE	ITT: second-line HCC patients	Regorafenib; placebo plus BSC			
MAIC using time-varying HRs (log-logistic model presented in CS)	CELESTIAL	Subpopulation: second-line HCC patients	Cabozantinib; placebo plus BSC	Anchored	Yes	Efficacy: OS; PFS Safety: N/a
	RESORCE	ITT: second-line HCC patients	Regorafenib; placebo plus BSC			
MAIC using independent parametric models (log-logistic or generalised gamma model presented in CS)[a] or OR (safety outcomes)	CELESTIAL	Subpopulation: second-line HCC patients	Cabozantinib	Unanchored	Yes	Efficacy: OS; PFS Safety: Diarrhoea; PPES
	RESORCE	ITT: second-line HCC patients	Regorafenib			

BSC - best supportive care; HCC - hepatocellular carcinoma; HR - hazard ratio; ITC - indirect treatment comparison; ITT, intention-to-treat; MAIC - matching-adjusted indirect comparison; OR - odds ratio; OS - overall survival; PFS - progression-free survival; PH - proportional hazards; AST - aspartate aminotransferase; PPES - palmar-plantar erythrodysaesthesia syndrome; CS - company's submission; N/a - not applicable

Notes: a – the company's clarification response (question A15) highlights that this model was incorrectly labelled as the generalised gamma model for OS in the CS

** Additional analysis presented as part of company's clarification response (question A13)*

3.3.2 Methods of company's ITC analyses

3.3.2.1 Bucher approach

The first ITC approach presented by the company includes a series of Bucher indirect comparisons. This form of comparison used aggregate-level data from the CELESTIAL and RESORCE trials,[15, 17] with placebo plus BSC as the common comparator arm. The relative treatment effect of cabozantinib versus regorafenib was estimated for efficacy outcomes including: OS, PFS, and three safety outcomes: hypertension, increased AST and fatigue. The CS[1] reports results from Bucher ITCs which utilised the ITT population from both CELESTIAL and RESORCE, where the CELESTIAL ITT population was broader than the RESORCE trial population as it included both second- and third-line patients.

The results of the Bucher ITCs are presented in Section B.3.10.2 of the CS.[1] The comparison for PFS was based on RECIST 1.1 criteria from the CELESTIAL trial[7] and modified RECIST (mRECIST) criteria from the RESORCE trial,[15] and the comparison for OS was based on a 2016 data-cut for the RESORCE trial. As part of their clarification response[19] (questions A13, A25 and A26), the company provided results from Bucher ITCs for both PFS and OS using the second-line subpopulation from CELESTIAL in order to more closely align with the RESORCE ITT population. Of note, this comparison using the second-line population of the CELESTIAL trial was based on RECIST 1.1 criteria for PFS for both trials, as well as the latest data-cut (2018) of the RESORCE trial for OS. Results from all Bucher ITCs are summarised in Section 3.3.3 of this report (see Table 11).

3.3.2.2 MAIC approach

The company also conducted a series of MAICs comparing cabozantinib versus regorafenib, citing differences in baseline characteristics between the CELESTIAL and RESORCE trials[7, 15] as a rationale for performing this type of ITC. These differences included the proportion of patients with ECOG PS 0, ethnicity and geographical regions (CS,[1] Section B.3.10.3). IPD were available for the CELESTIAL trial. In the MAIC analyses, the company used a subpopulation of the CELESTIAL trial, specifically second-line HCC patients who had prior treatment with sorafenib (i.e., "pure" second-line patients) to better align the population with that of the RESORCE trial (which only evaluated second-line patients). The company's clarification response[19] (question A14) provides details of the sample size of the second-line population with complete baseline characteristics (i.e., after exclusion of subjects with missing covariate data): a total of 484 patients were included in the MAIC analysis (out of a total of 707 patients included in the ITT trial population). The ERG notes that no attempt was made by the company to impute missing covariate data in the CELESTIAL trial.

Aggregate-level baseline characteristics and outcome data were extracted for the RESORCE trial;[15] Kaplan-Meier curves for PFS and OS were digitised and pseudo IPD were reconstructed using the algorithm reported by Guyot *et al.* (2012).[21] The proportion of patients experiencing individual AEs

was also extracted, including: increased AST; elevated bilirubin; fatigue; hypertension; diarrhoea and PPES. During the clarification stage, the ERG asked the company to present results from a MAIC evaluating all Grade 3/4 AEs combined, rather than individually (see clarification response,[19] question A19). However, the company did not undertake this analysis as they stated that the incidence of Grade 3/4 AEs was almost identical between the two treatment arms. However, without quantifying the results for this analysis, there remains uncertainty around the treatment effect for this outcome.

The CS[1] (Section 3.10.3) states that the baseline characteristics which were used to inform the weighting process were selected from the preliminary set based on their potential influence on key efficacy outcomes (PFS and OS) and AEs. Baseline characteristics for the company's base case analyses (denoted Scenario "S1") were justified for inclusion in the MAIC based on feedback received by clinical experts from a UK advisory board meeting and were further confirmed at a second advisory board meeting. The potential effect modifiers included: age; race; geographical region; ECOG PS; Child-Pugh grade; duration of prior sorafenib treatment; extrahepatic disease; macrovascular invasion; aetiology of HCC (Hepatitis B, alcohol use and Hepatitis C) and alpha fetoprotein level (AFP) level. The company also explored a second scenario (denoted Scenario "S2") which included only effect modifiers for OS (primary survival outcome), identified using a stepwise regression approach. The subset of factors included: gender; ECOG PS; extrahepatic disease; macrovascular invasion and AFP level; however, the CS states that following clinical feedback received from the advisory board, gender was not included in the matching. In response to a request for clarification from the ERG[19] (question A18), the company provided further information around the selection of these factors, and clarified that the classification of each factor was determined by how the data were reported in the RESORCE trial;[15] all factors included in the matching (apart from duration of prior sorafenib and aetiology of disease) were reported as dichotomous variables. Duration of sorafenib was retained as a continuous variable and aetiology of disease was split into multiple categories. A summary of the effect modifiers and their respective classification is presented in Table 10.

Table 10: Summary of effect modifiers included in company's matching (adapted from clarification response, question A18)

Factor included in matching	Classification of factor	Factor included in Scenario 1 (S1)	Factor included in Scenario 2 (S2)	Rationale for classification
Age	< 65 years ≥ 65 years	✓		To reflect average age in RESORCE. This was categorised to minimise impact on ESS
Race	Female Male	✓		Binary variable
Geographical region	Asia Other	✓		To reflect reporting of RESORCE trial region baseline characteristic
ECOG performance status	ECOG 0 ECOG 1 or 2	✓	✓	Binary variable. ECOG 1 and ECOG 2 combined due to low ECOG 2 numbers
Child Pugh grade	A, B or unknown	✓		Binary variable
Duration of prior sorafenib	Continuous variable	✓		-
Extrahepatic disease	Present Absent	✓	✓	Binary variable
Macrovascular invasion	Present Absent Unknown	✓	✓	Binary variable
Aetiology of HCC (Hepatitis B)	Present Absent Unknown	✓		Binary variable
Aetiology of HCC (Alcohol use)	Present Absent Unknown	✓		Binary variable
Aetiology of HCC (Hepatitis C)	Present Absent Unknown	✓		Binary variable
AFP level	≥ 400 ng/ml < 400 ng/ml	✓	✓	To reflect reporting of RESORCE trial AFP level baseline characteristic. This is a diagnostic threshold for HCC

ESS - effective sample size; AFP - alpha fetoprotein level; ECOG - Eastern Cooperative Oncology Group; HCC - hepatocellular carcinoma

The ERG's clinical advisors considered that Scenario S1 included the key prognostic variables and treatment effect modifiers; however, they also suggested that the number of prior local regional therapies and Barcelona Clinic Liver Cancer (BCLC) staging were additional important prognostic factors. The ERG notes that there is a potential for the presence of strong correlation between BCLC stage and Child Pugh grade and ECOG PS if these variables are considered to measure similar aspects of health, which may result in overmatched data and an unnecessary reduction in the effective sample

size (ESS) if included in the matching process. Further, BCLC staging was not captured in the CELESTIAL trial[7] and could not be matched on. The ERG's clinical advisors suggested that Child Pugh grade, extrahepatic disease and ECOG PS were considered as being particularly important potential prognostic factors and/or treatment effect modifiers.

In response to a request for clarification from the ERG[19] (question A23), the company provided histograms which display the distribution of estimated weights obtained for Scenarios S1 and S2. These are provided in Figures 14 and 15 of the company's clarification response. The histograms indicate that some individuals were assigned large weights in S1, with a maximum rescaled weight of 9.21. In S2, no extreme weights were observed, with a maximum rescaled weight of 1.61. However, the company acknowledges that S2 does not include matching on some baseline characteristics which were identified as being important effect modifiers. In their response to clarification question A21, the company confirmed that the weights from S1 were used in the company's anchored and unanchored MAICs conducted for OS and PFS; no results were presented for these outcomes using weights from S2.

In S1, the ESS for the cabozantinib arm was reported by the company to be 187.27 (57.4% of the original sample size [N=326]). A small ESS indicates that weights are highly variable due to a lack of population overlap and that the resulting estimates may be unstable.[22] Whilst the ESS was notably higher for S2 (ESS=303.24), this has been at the expense of matching on fewer effect modifiers in an attempt to balance trial populations. The ERG notes that there may be residual confounding if all effect modifiers are not accounted for in the matching process.[22]

Three types of population-adjusted analyses were performed for PFS and OS using the weights from Scenario S1, including:

- Anchored comparisons which apply the PH assumption, and which utilise a constant HR estimated from a Cox regression model using weighted CELESTIAL data and RESORCE data to provide a comparison for cabozantinib versus regorafenib.
- Anchored comparisons which do not assume PH, and which explore if there is any difference in the treatment effect emerging between cabozantinib and regorafenib over time by generating time-varying HRs from hazard profiles of fitted parametric models to the weighted CELESTIAL data and RESORCE data.
- Unanchored comparisons, which evaluate absolute effects through fitting parametric models to weighted cabozantinib data and regorafenib data.

One form of anchored MAIC conducted by the company was based on the estimation of a constant HR to represent the treatment effect between cabozantinib and regorafenib. The company also provided

results from a time-varying anchored MAIC analysis, fitting parametric models to the weighted cabozantinib and regorafenib arms of their respective trials;[7, 15] the company stated that that a log-logistic model was considered to provide the best fit for both OS and PFS based on an assessment of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics (CS,[1] Section B.3.10.3). At the clarification stage, further details were requested from the company regarding the time-varying approach, including the methodology adopted to estimate the time-varying HR after fitting parametric curves to the data, and further, how the weights from the matching were incorporated into the ITC (see clarification response,[19] question A20). The company's response states that independent parametric models of the same distribution were fitted to both arms of the weighted CELESTIAL data and both arms of the RESORCE trial to generate a hazard function for each treatment arm. The time-varying HR between cabozantinib versus placebo from the CELESTIAL trial was generated by dividing the hazard for the cabozantinib parametric model by the hazard for the placebo parametric model at each timepoint. The time-varying HR for regorafenib versus placebo from the RESORCE trial was generated in a similar way. The time-varying HR for cabozantinib versus regorafenib was then estimated by calculating the ratio of the cabozantinib versus placebo HR versus the regorafenib versus placebo HR. Weights from the population-adjustment process were incorporated into the analysis by fitting weighted parametric survival models.

An unanchored MAIC analysis was also performed by the company, which was undertaken by fitting a series of independent parametric models to the weighted cabozantinib and regorafenib arms of the respective trials.[7, 15] The CS[1] states that a generalised gamma distribution was considered to provide the best fit to the data based on an assessment of AIC and BIC statistics. However, the company's clarification response[19] (question A15) includes a correction which states that the best fitting model for OS should have been labelled as the log-logistic model.

Results from all MAIC analyses are presented in Section B.3.10.3 of the CS.[1] The results for comparisons of efficacy and safety are summarised in Table 11 and Table 12, respectively (see Section 3.3.3).

3.3.3 Results of company's ITC analyses

The results of the ITCs presented in the CS[1] are summarised in Table 11 (efficacy outcomes, including OS and PFS) and Table 12 (safety outcomes, including hypertension, elevated AST, fatigue, elevated bilirubin, diarrhoea and PPES).

Table 11: Summary of company's ITC analyses conducted for efficacy outcomes

Line of treatment	Analysis	Description	Efficacy outcomes, HR (95% CI)		
			Comparison	OS	PFS
Second 72%, third 28%	Bucher	ITC anchored on placebo plus BSC, without adjustment for cross-trial differences	Cabozantinib vs. regorafenib	1.21 (0.90, 1.62)	0.96 (0.73, 1.26)
Second	Bucher	ITC anchored on placebo plus BSC, without adjustment for cross-trial differences		1.13[a,c] (0.83, 1.53)	0.93[a,b] (0.69, 1.25)
	Anchored MAIC (Constant HR)[d,e]	Weighted Cox PH regression model (where weights are estimated from matching on trial baseline characteristics)		1.15 (0.79, 1.69)	0.79 (0.56, 1.11)
	Anchored MAIC (Time-varying HR)	The company selected a log-logistic model as the best fitting model to estimate a time-varying HR for both OS and PFS		Time-varying HR>1.0, suggesting a trend of improved OS for regorafenib over cabozantinib. Results across the models show that over time, the HR is not statistically different from 1.0 (95% CI interval includes a time-varying HR of 1.0)	Time-varying HR<1.0, suggesting a trend of improved PFS for cabozantinib over regorafenib. Results across the models show that over time, the HR is not statistically different from 1.0 (95% CI interval includes a time-varying HR of 1.0)
	Unanchored MAIC	The company selected a log-logistic model[d] (OS) and generalised gamma model (PFS) fitted to weighted cabozantinib and regorafenib arms		Large amount of overlap until year 1 when cabozantinib begins to show a relatively small benefit over regorafenib. Cabozantinib has a larger point estimate for mean OS (24.65 vs. 21.17 months) and a higher median OS (11.40 versus 10.29 months).	Statistically significant benefit for cabozantinib until approximately 1 year when the PFS curves show minimal difference for the rest of the time horizon. Cabozantinib has a larger point estimate for mean PFS than regorafenib (7.17 vs. 6.04 months) and higher median PFS (5.49 vs. 3.39).

ITC - indirect treatment comparison; BSC - best supportive care; CI - confidence interval; HR - hazard ratio; MAIC - matching-adjusted indirect comparison; OR - odds ratio; OS - overall survival; PFS - progression-free survival; PH - proportional hazards

Notes: HR<1.0 favours cabozantinib over regorafenib, a - analysis conducted in response to clarification question A13, using second-line subgroup data from CELESTIAL trial; b - analysis conducted in response to clarification question A26, using RECIST 1.1 PFS data for both CELESTIAL and RESORCE trials (instead of using RECIST 1.1 in CELESTIAL and mRECIST in RESORCE); c - analysis conducted in response to clarification question A25, using data cut from the RESORCE trial from 2018 (instead of using data cut from the RESORCE trial from 2016); d - a correction has been made by the company which stated that the best fitting model for OS in the unanchored comparison was the log-logistic model instead of the generalised gamma model; e - Weibull model with a constant HR was also explored by the company as part of a response to clarification question B6

Table 12: Summary of company's ITC analyses conducted for safety outcomes

Line of treatment	Analysis	Description	Safety outcomes, OR (95% CI)						
			Comparison	**Hypertension**	**Elevated aspartate aminotransferase**	**Fatigue**	**Elevated bilirubin**	**Diarrhoea**	**PPES**
Second 72%, third 28%	Bucher	ITC anchored on placebo plus BSC, without adjustment for observed cross-trial differences	Regorafenib vs. cabozantinib[a]	0.2 (0.0-1.2)	0.6 (0.2-1.6)	1.2 (0.3-5.6)	-	-	-
Second	Anchored MAIC	Weighted OR (where weights are estimated from matching on trial baseline characteristics)[b]	Cabozantinib vs. regorafenib	8.17[c] (0.90, 73.70)	2.20[c] (0.63, 7.84)	1.09[c] (0.17, 6.96)	0.78[c] (0.07, 9.30)	-	-
	Unanchored MAIC		Cabozantinib vs. regorafenib	-	-	-	-	**5.70[c] (2.72, 11.94)**	1.05[c] (0.67, 1.65)

ITC - indirect treatment comparison; BSC - best supportive care; CI - confidence interval; HR - hazard ratio; MAIC - matching-adjusted indirect comparison; OR - odds ratio; OS - overall survival; PFS - progression-free survival; PPES - palmar-plantar erythrodysaesthesia syndrome

Notes: Bucher: OR>1 favours cabozantinib over regorafenib; MAIC: OR<1 favour cabozantinib over regorafenib; bold denotes statistically significant comparison at 5% level; a - the ERG believes this comparison to be incorrectly labelled as cabozantinib versus regorafenib in both the CS and the response to clarification question A16; b - results based on weights obtained from Scenario S1; c - results transformed by the ERG from logOR to OR using the exponential function

3.3.3.1 Bucher approach

A summary of the results from the Bucher ITCs for OS and PFS are presented in the CS[1] (Table 24) and are also summarised in Table 11. The CS states that the results from the Bucher ITCs showed HR estimates which favoured cabozantinib over regorafenib for PFS (HR<1.0) and which favoured regorafenib over cabozantinib for OS (HR>1.0), but the results were statistically non-significant, which the company suggests reflects similar efficacy in terms of OS and PFS for both treatments. For the Bucher ITC analysis using the second-line subpopulation from CELESTIAL[7] (presented in response to clarification question A13[19]), the company also suggested that there was no statistically significant difference between the two treatments in this subgroup of patients and therefore, the conclusions remain unchanged.

A summary of the results from the Bucher ITCs for safety outcomes are presented for cabozantinib versus regorafenib in Table 26 of the CS;[1] these are also summarised in Table 12. The CS states that a Bucher ITC was only feasible when there were events in all arms of the CELESTIAL and RESORCE trials.[7, 15] Therefore, only three AEs were compared: hypertension, elevated AST and fatigue. The CS states that the results show no statistically significant differences between the AE OR estimates for cabozantinib and regorafenib. Further, the CS (Section B.3.10.4) states that the ITC results suggest that cabozantinib has *"similar tolerability compared to regorafenib.*" However, the OR point estimates from the Bucher ITCs conducted for hypertension and elevated AST differ from 1.0. The company's response to clarification question A24[19] regarding the assumption of similar tolerability between cabozantinib and regorafenib suggests that since the *p*-value for hypertension and AST is not significant, and the clinical experts advising the company agreed that the safety profiles of cabozantinib and regorafenib are similar, this may indicate that the tolerability of both regimens is considered to be the same. However, the company's clarification response (question A16) suggests that the Bucher comparisons presented for safety outcomes (hypertension, elevated AST and fatigue) are incorrectly labelled. Upon clarification, the ERG believes the company has also mislabelled the treatment effect in Table 7 of the clarification question response document for both CELESTIAL and RESORCE, which in fact, represent the treatment effect between placebo versus cabozantinib and placebo versus regorafenib, respectively. The ERG has re-labelled the OR estimates from a Bucher ITC for three safety outcomes as a comparison between regorafenib versus cabozantinib (instead of cabozantinib versus regorafenib); these results are presented in Table 12.

The ERG believes that the Bucher ITC approach adopted by the company does not provide robust estimates of comparative efficacy and safety due to the presence of observed cross-trial differences. In addition, the results from the ITCs presented in Table 24 of the CS[1] are further limited by the fact that CELESTIAL data[7] reflect the ITT population and do not utilise data from the second-line subpopulation from the trial. The company's clarification response[19] (question A13) provides estimates of the Bucher

ITCs for PFS and OS using the second-line population from the CELESTIAL trial. Whilst the conclusions of this analysis remain unchanged from those presented for the CELESTIAL ITT population, the results of this analysis may not be reliable due to the remaining cross-trial differences between the CELESTIAL and RESORCE trial populations.

3.3.3.2 MAIC approach

Anchored comparisons

The CS[1] provides the results of MAIC analyses utilising a subpopulation from the CELESTIAL ITT population,[7] specifically second-line HCC patients who had prior treatment with sorafenib (i.e., pure second-line patients). The ERG agrees that this initial equalisation of study populations is an appropriate step prior to conducting an ITC. The results of the anchored comparison for efficacy outcomes (OS and PFS) between cabozantinib and regorafenib using a constant HR estimated from a Cox regression model are shown in Table 31 of the CS; these are also summarised in Table 11. The point estimate of the HR for cabozantinib versus regorafenib favours PFS cabozantinib, whilst the point estimate of the HR for OS favours regorafenib. Both of these results are statistically non-significant (95% CIs include an HR estimate of 1.0), from which the company concludes that there is no evidence of a difference between the treatments. The ERG believes that the MAIC analysis using a constant HR have been performed appropriately.

During the clarification stage, the ERG asked the company to provide the unweighted and weighted Kaplan-Meier curves for the cabozantinib arm from CELESTIAL[7] (see clarification response,[19] question A15). These are reproduced in Figure 4 and Figure 5 for PFS and OS, respectively, using weights from both Scenarios S1 and S2. The Kaplan-Meier curves show the PFS and OS data prior to-(unweighted) and post-adjustment (weighted), using the weights obtained from the matching process. The company concludes that Scenarios S1 and S2 yield similar results. However, relative to the unweighted curve, the use of weights from Scenario S1 results in a greater shift in the Kaplan-Meier curve compared to the weights from S2, and this trend is observed for both PFS and OS. This is an expected result given the greater reduction in ESS when using weights from Scenario S1 compared to S2.

Figure 4: Unadjusted and weighted Kaplan-Meier curves for PFS, cabozantinib arm of CELESTIAL, (reproduced from clarification response, question A15)

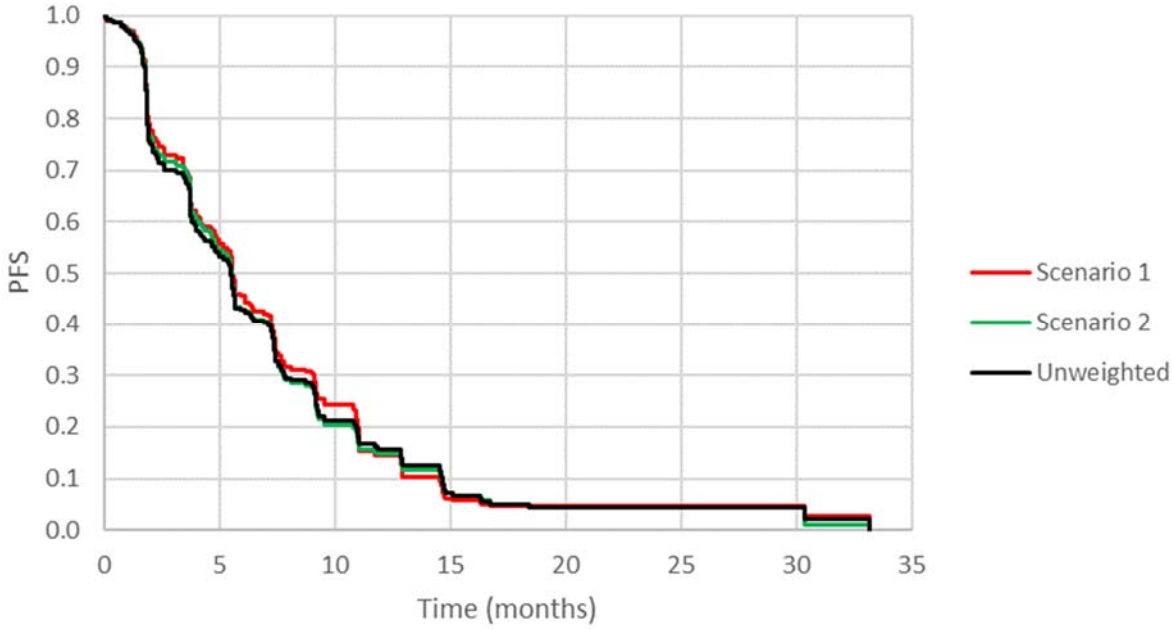
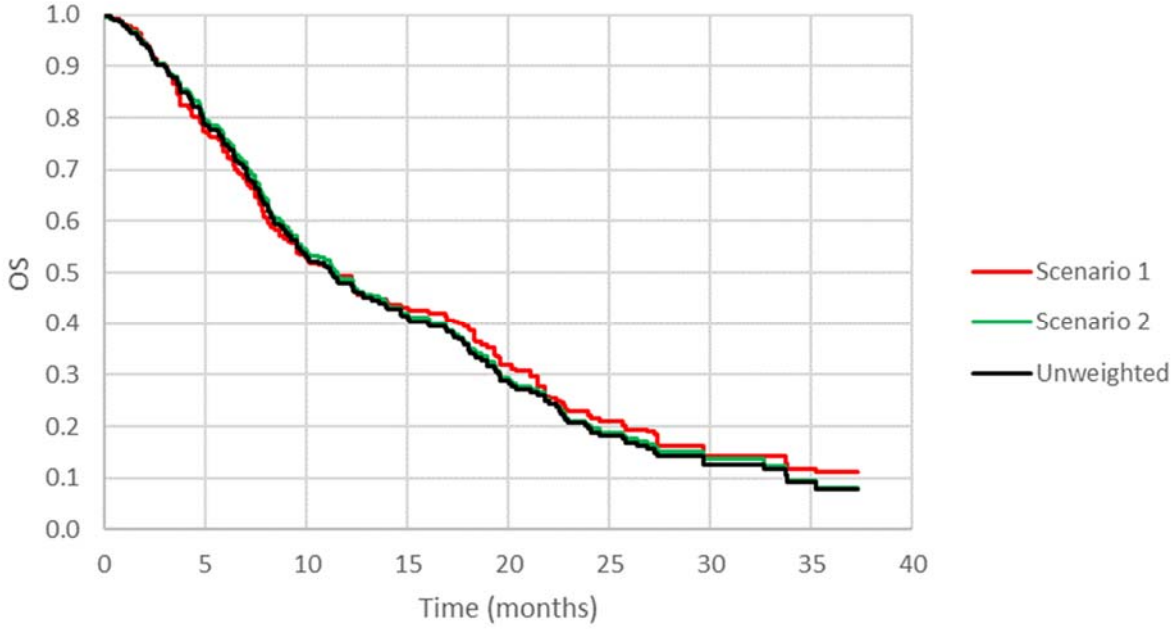


Figure 5: Unadjusted and weighted Kaplan-Meier curves for OS, cabozantinib arm of CELESTIAL, (reproduced from clarification response, question A15)



The company also used the MAIC methodology to evaluate six AE outcomes. An anchored approach was adopted for the analysis of four AEs: increased AST; elevated bilirubin; fatigue and hypertension. An unanchored approach was adopted for the analysis of diarrhoea and PPES due to zero event rates in the placebo arms of the RESORCE and CELESTIAL trials,[7, 15] respectively. The treatment effect for each AE was represented by a log odds ratio (LOR) and associated 95% CI. Results are presented in Table 30 of the CS;[1] these results are also summarised Table 12. The incidence of AEs was generally

higher for cabozantinib than regorafenib; however, a statistically significant difference was only observed for diarrhoea, and this was based on an unanchored comparison approach.

For the analysis of PFS and OS, the company explored the PH assumption using weighted second-line cabozantinib data from Scenario S1[7] and regorafenib data from the RESORCE trial.[15] The CS[1] states that the PH assumption was not satisfied for PFS or OS based on an assessment of the log cumulative hazard plots, Schoenfeld residuals and the Grambsch-Therneau test. The CS states that due to the uncertainty around the PH assumption, an alternative time-varying HR analysis was performed. The company conducted an anchored analysis, based on the assumption that the PH assumption did not hold, to explore any differences in the treatment effect emerging between cabozantinib and regorafenib over time. The CS states that a log-logistic model was selected for the time-varying approach as it was the best fitting model according to the AIC and BIC statistics. The results of the anchored analysis using time-varying HRs generated from the log-logistic model are presented in the CS (Section 3.10.3, Figures 17-18) for PFS and OS; the key findings are summarised in Table 11. The CS states that other parametric models were tested using a time-varying approach, including Weibull, Gompertz, log-normal and generalised gamma distributions; these are presented in CS Appendix I. The company's clarification response[19] (question A20) provides further information regarding the approach adopted to estimate a time-varying HR. The ERG believes that the time-varying approach adopted by the company has been undertaken appropriately.

Unanchored comparisons

The results of the unanchored analysis for PFS and OS are shown in CS[1] Figures 19 and 20; these are also summarised in Table 11. The company's clarification response[19] (question A15) provides the AIC and BIC statistics for each of the models fitted in the unanchored comparison, which the company used to support the selection of the log-logistic model (this model provided the lowest AIC and BIC values for the weighted cabozantinib arm). However, results for other parametric models (i.e. those explored as part of the anchored comparisons) were not presented by the company.

The unanchored approach using a generalised gamma model for PFS showed a statistically significant benefit for cabozantinib until approximately 1 year; beyond this timepoint the PFS curves show little difference for the remainder of the time horizon. Cabozantinib had a larger point estimate for mean PFS than regorafenib and a higher median PFS. The OS curves based on the log-logistic model showed a large amount of overlap until year 1 when cabozantinib begins to show a relatively small benefit over regorafenib. Cabozantinib had a larger point estimate for mean OS and a higher median OS. The company concluded that the results across the models show that over time, the HR is not statistically different from 1.0, suggesting no difference in treatment effect. Furthermore, the HR is generally seen to be constant and near 1.0 as the treatment effect is extrapolated, which suggested equivalence in

treatment effect over time. The ERG has concerns that the direction of the treatment effect for OS is not consistent across the different ITC analyses presented by the company - both the Bucher ITC and anchored MAICs (constant HR and time-varying HR) yield HRs which are greater than 1.0 for cabozantinib versus regorafenib (favouring regorafenib), whereas the results from the unanchored MAIC suggests an OS benefit for cabozantinib over regorafenib.

The company's clarification response[19] (question A15) provides other fitted parametric models overlaid on the cabozantinib Kaplan-Meier curves, as shown in Figure 6 (PFS) and Figure 7 (OS). For OS, the log-logistic model does not appear to provide a good fit to the tail of the data and provides the most optimistic estimates of long-term survival (along with the log-normal model).

Figure 6: Parametric curves overlaid on top of the weighted cabozantinib Kaplan-Meier curve for PFS (reproduced from clarification response, question A15)

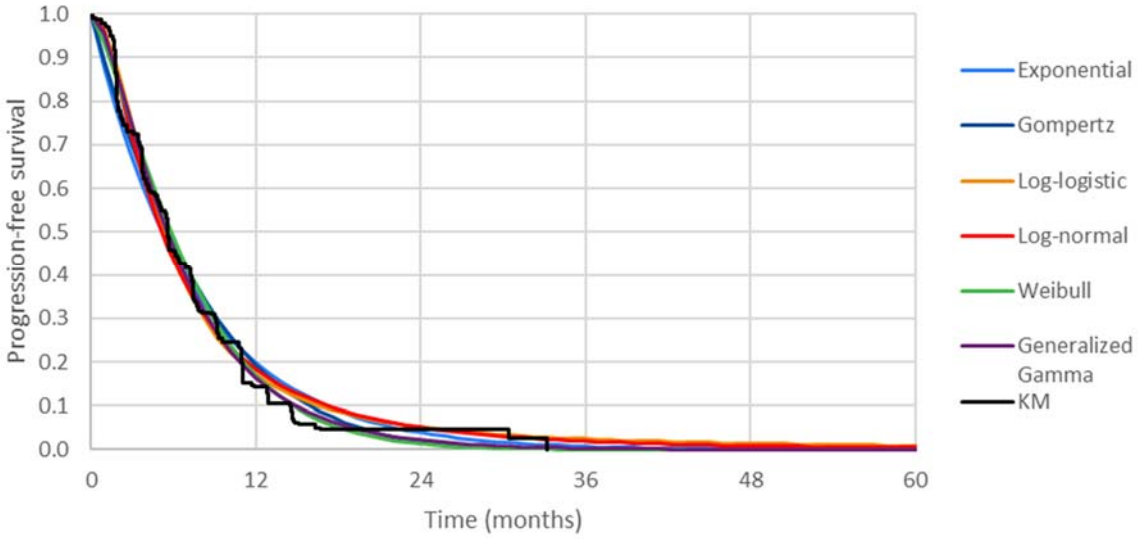
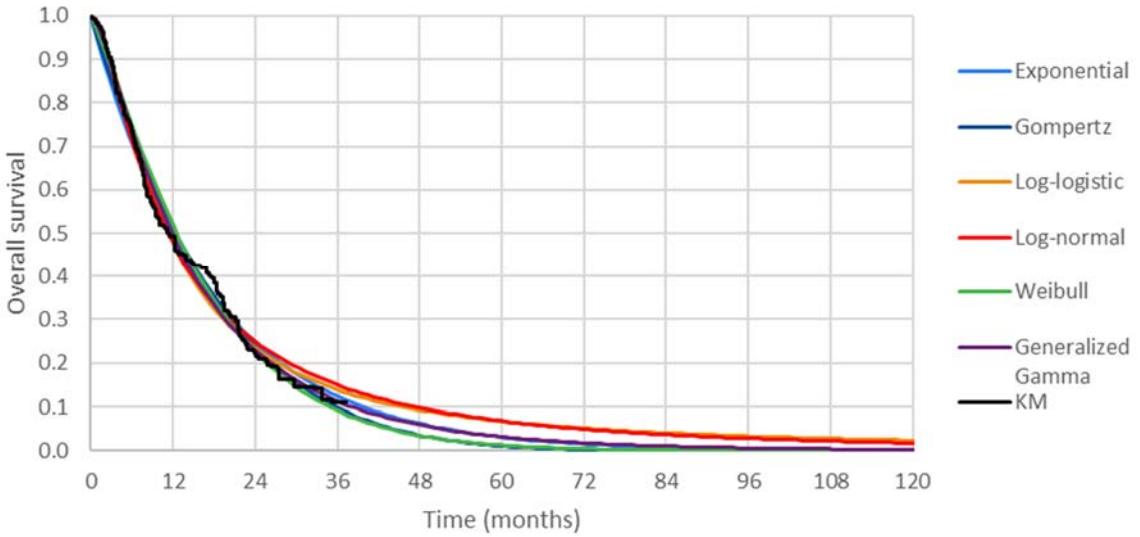


Figure 7: Parametric curves overlaid on top of the weighted cabozantinib Kaplan-Meier curve for OS (reproduced from clarification response, question A15)



During the clarification process, the ERG asked the company to provide the empirical and smoothed hazard functions (see clarification response,[19] question A15). The company's response provided this for the cabozantinib arm of the CELESTIAL trial[7] and the regorafenib arm of the RESORCE trial[15] in Figures 8 and 9 for OS and Figures 12 and 13 for PFS of the clarification response document, along with the hazard function of the log-logistic and generalised gamma models overlaid (best fitting models for OS and PFS, respectively). The shape of the smoothed hazard function does not follow the same shape as the hazard function of the log-logistic model (selected for OS), which suggests that this may not be the most appropriate model selection. However, for PFS, the smoothed hazard function follows a similar shape to the hazard function of the generalised gamma model, which suggests that this may be an appropriate model selection.

The ERG believes that the unanchored comparisons presented by the company may not be reliable; this form of comparison relies upon strong assumptions which are rarely satisfied, for example, matching on all prognostic factors and treatment effect modifiers, and relies on a comparison of absolute effects, which does not preserve trial randomisation. However, the ERG also recognises that the placebo plus BSC arm of both CELESTIAL and RESORCE trials[7, 15] may differ: the company has shown that the HR between the two placebo plus BSC arms is not equal to 1.0, which suggests that the anchor arm may not be entirely comparable between the two trials. This is a notable limitation of the anchored comparisons, which rely on the assumption of transitivity (i.e. the anchor arm is comparable between the two trials).

The company's clarification response[19] (questions A12 and B6) confirms that the anchored MAIC analysis using a constant HR from the Weibull model is considered to be their base case. This ITC is also denoted as the company's base case scenario in their analysis of incremental quality-adjusted life years (QALYs) gained for cabozantinib versus regorafenib (see Section 4.5.1).

3.3.4 Key limitations of company's ITC analyses

The ERG believes that the company's ITC analyses are subject to a number of uncertainties. Whilst preserving trial randomisation, the use of the Bucher ITC approach is limited by the lack of adjustment for cross-trial differences which are present in the data. One of the key assumptions underpinning the Bucher approach is that there is no difference between trials regarding the distribution of effect modifiers. The company acknowledges that there are considerable observed differences in trial populations (CS,[1] Section 3.10.3, page 61), a consequence being that this assumption is unlikely to be satisfied. Further, the full ITT population from the CELESTIAL trial[7] was used in the Bucher comparison presented in the CS,[1] meaning that second- and third-line patients were compared against second-line patients from the RESORCE trial.[15] Therefore, results from this form of comparison are unlikely to be sufficiently robust to draw inferences. Despite the company also presenting results using

the second-line population from the CELESTIAL trial, the ERG believes that this form of comparison may lack robustness due to the remaining cross-trial differences between the studies.

The company has conducted a series of population-adjusted ITCs in attempt to overcome cross-trial differences between the CELESTIAL and RESORCE trials.[7, 15] Despite the company utilising the subpopulation of second-line HCC patients who had prior treatment with sorafenib patients from the CELESTIAL trial to align more closely with the population from the RESORCE trial, the ERG has concerns that there may remain differences between the two trials which have not been accounted for in the ITC analyses. Further, whilst anchored comparisons are a recommended approach for ITCs, as they provide a way of comparing two interventions with fewer assumptions required than unanchored comparisons, it is important that the anchor arm (in this case, placebo plus BSC) is consistent across both trials. There are concerns with regard to the comparability of the placebo plus BSC arms across both CELESTIAL and RESORCE trials. Specifically, the company evaluated the treatment effect between the placebo arms of both trials and found that the HR for OS was different from 1.0, although this result was statistically non-significant (HR=0.87; 95% CI [0.67-1.15]; *p*=0.326). For PFS, there was a greater difference between the two trials (HR=0.69; 95% CI [0.55-0.87]; *p*=0.002). A similar result was found for both OS and PFS in Scenario S2. The assumption of transitivity which underpins anchored ITCs may be violated if there are systematic differences in the anchor arm of each trial. The company acknowledges that this finding suggests that there may remain important cross-trial differences which have not been addressed in the MAIC, raising concerns on the robustness of the anchored ITC analyses conducted.

Identification of the baseline characteristics included in the matching process was based on clinical input; however, in Scenario S2, the factors were selected using stepwise regression methods; an empirical approach informed by the data. The ESS estimate for the cabozantinib arm in Scenario S1 is approximately 54% of the original sample size after weighting, showing a substantial reduction in the number of patients informing the analysis. The ESS estimate for the cabozantinib arm was higher in S2, being approximately 93% of the original sample size; however, fewer factors were included in the matching process meaning that important effect modifiers may not have been accounted for and therefore, residual confounding may be present.[22]

The company has also performed an unanchored comparison, comparing cabozantinib versus regorafenib without utilising data from the placebo plus BSC arm from either trial. Unanchored comparisons do not preserve trial randomisation and are limited by the comparison of absolute effects only. Further, unanchored comparisons rely on strong assumptions - that all prognostic factors and treatment effect modifiers are accounted for in the matching process. This condition is rarely met.[22] Therefore, the company's unanchored comparisons may not be robust if there are other factors

considered influential on outcomes or which may alter the treatment effect between cabozantinib and regorafenib. Remaining differences between study populations may result in the presence of residual confounding, meaning that the unanchored MAIC results are limited. Further, the findings from the unanchored comparison conducted for OS (which uses a log-logistic model fitted to both treatment arms) are inconsistent with those obtained from the Bucher and anchored MAIC analyses, where cabozantinib was found to be superior to regorafenib (higher mean and median OS), despite the HR estimates from the anchored comparisons being greater than 1.0. This inconsistent finding may suggest uncertainty around the treatment effect, but it may be an artefact of comparing absolute effects and breaking trial randomisation. Therefore, the results from the unanchored comparison may not be reliable.

The results of the ITC analyses presented by the company were used to justify an assumption of equivalent clinical effectiveness between cabozantinib and regorafenib; this assumption underpins the company's cost-comparison analysis (see Section 4). However, whilst a statistically non-significant difference has been found between cabozantinib and regorafenib, this does not infer equivalence of the two regimens. The ERG believes that the Bucher ITCs performed by the company are limited because they do not account for cross-trial differences which have been identified. The unanchored ITC is also limited by lack of preservation of trial randomisation and the potential problem of residual confounding. The ERG considers the anchored MAIC analyses to provide the most robust estimates of relative treatment effects between cabozantinib and regorafenib; however, like the company, the ERG also has concerns regarding the comparability of the anchor arm (placebo plus BSC) across the two trials. The analysis conducted by relaxing the PH assumption by allowing the HR to vary with time may be the most appropriate approach in light of violation of the PH assumption for PFS and after assessment of the time-varying HR plots, which show that the HR is not constant for a number of parametric models.

3.3.5 Conclusions on the company's ITCs

The company has explored a number of statistical ITC approaches, all of which show a statistically non-significant difference between cabozantinib and regorafenib. The ERG believes that there are limitations associated with all ITC analyses conducted; however, an anchored approach would be the preferred form of ITC to estimate comparative efficacy and safety in the absence of direct head-to-head data. Due to the limitations and concerns outlined, the ERG believes that there remains uncertainty around the treatment effect between cabozantinib and regorafenib, including the assumption of equivalence of the two regimens and therefore, the results of the ITCs conducted should be interpreted with caution.

4. ERG'S CRITIQUE OF THE COMPANY'S COST COMPARISON ANALYSIS

4.1 Model summary, assumptions and evidence sources

As part of the CS,[1] the company submitted an executable cost comparison model of cabozantinib versus regorafenib for patients with previously treated advanced HCC. The company's base case analysis estimates the cost savings for cabozantinib versus regorafenib based on the number of whole packs of cabozantinib or regorafenib required to treat patients until progression and the cost per pack of each drug, assuming the same treatment duration for both groups. The model applies a 15 year time horizon to estimate maximum treatment duration based on PFS data from the CELESTIAL trial ITT population[7] as a proxy. Discounting is not included. All analyses presented in the CS use point estimates of parameters; probabilistic analysis has not been undertaken. The company's analyses include the Patient Access Scheme (PAS) discount for cabozantinib and the list price for regorafenib. The results of the company's analyses including both the PAS price for cabozantinib and the comparator Patient Access Scheme (cPAS) price for regorafenib are provided in a separate confidential appendix to this ERG report.

The company's base case analysis makes the following assumptions:

(i) *Equivalent clinical outcomes*. Cabozantinib and regorafenib are assumed to be clinically equivalent in terms of PFS, OS, time on treatment (ToT) and AEs. As such, the incremental QALY gain for cabozantinib versus regorafenib is assumed to be zero.

(ii) *Treatment is given until disease progression.* Whilst the RESORCE and CELESTIAL trials[7, 15] permitted some patients to continue treatment beyond disease progression, the cost comparison model assumes that both drugs are given until disease progression in all patients. PFS duration is estimated using a log-normal model fitted to IPD from the cabozantinib arm of the CELESTIAL trial.[7] The executable model does not include the cumulative probabilities of PFS over time; rather, all calculations are undertaken using the 15-year restricted mean (i.e., the area under the curve [AUC] up to 15 years after starting treatment). The impact of an arbitrary increase/decrease in mean treatment duration for both groups (+/-20%) is tested in sensitivity analysis.

(iii) *No difference in costs except for drug acquisition.* The only difference in costs between the treatment groups in the base case analysis relates to the costs of drug acquisition. The model assumes that there is no difference in the costs of disease management (e.g., clinic visits and monitoring), subsequent anticancer therapies given after disease progression or AEs between the treatment groups. Both drugs are given orally; hence, administration costs are not relevant. The impact of applying treatment-specific AEs on costs is tested in sensitivity analysis.

(iv) *Perfect relative dose intensity.* The drug acquisition cost calculations assume 100% relative dose intensity (RDI) in both groups (i.e., patients receive the full recommended dose on every day that they receive either drug, irrespective of whether their dose has been reduced). This assumption is tested in sensitivity analysis.

(v) *Wastage costs included.* Both cabozantinib and regorafenib are subject to additional costs associated with wastage; these are captured by estimating the number of full packs of treatment required to treat patients up to the mean PFS duration. The effect of removing this assumption (by splitting packs) is tested in sensitivity analysis.

The company's base case and sensitivity analyses are summarised in Table 13. The evidence sources used to inform the company's model are summarised in Table 14. The frequencies of AEs and associated management costs, as applied in the sensitivity analysis, are summarised in Table 15.

Table 13: Summary of cost comparison analyses presented in the CS

Analysis	Description of analysis
Base case	Assumes equivalence in PFS, OS, ToT and AEs. Includes wastage costs (number of full packs required). Excludes AE costs. Assumes perfect RDI.
SA1 - Time on treatment – 20% ([redacted] months)	Same as base case analysis, but assumes ToT for both drugs is 80% of the mean value
SA2 - Time on treatment + 20% ([redacted] months)	Same as base case analysis, but assumes ToT for both drugs is 120% of the mean value
SA3 - Include drug-specific AE costs	Same as base case analysis, but includes AE frequencies for cabozantinib and regorafenib using RESORCE[15] and the company's MAICs to estimate cost differences
SA4 - Include RDI	Same as base case analysis, but uses mean daily dose received in RESORCE[15] and CELESTIAL[7] to estimate number of packs required
SA5 - Exclude wastage costs	Same as base case analysis, but assumes that packs can be split

SA - sensitivity analysis; PFS - progression-free survival; OS - overall survival; ToT - time on treatment; AE - adverse event; RDI - relative dose intensity; MAIC - matching-adjusted indirect comparison

Table 14: Evidence sources used to inform the company's cost comparison model

Parameter	Value	Source	ERG comments
Mean time on treatment – both treatment groups	[redacted] months	Log-normal model fitted to PFS data from CELESTIAL[7]	The model estimates that [redacted] packs of cabozantinib and regorafenib are required to treat to progression
Cost per pack – cabozantinib	List price: £5,143.00 PAS price: [redacted]	BNF[12]	Pack size is 30 x 60mg tablets (30 days' supply)
Cost per pack – regorafenib	List price: £3,744 cPAS price: see confidential appendix	BNF[12]	Pack size is 84 x 40mg tablets (28 days' supply)
RDI – cabozantinib (SA only)	0.61	CELESTIAL[7]	Calculated from mean vs. planned dose in trial. Base case analysis assumes RDI=100%
RDI – regorafenib (SA only)	0.90	RESORCE[15]	Calculated from mean vs. planned dose in trial. Base case analysis assumes RDI=100%
AE frequency – cabozantinib (SA only)	See Table 15	MAIC using data from RESORCE and CELESTIAL[1]	Calculated from ORs presented in CS[1] Table 35
AE frequency – regorafenib (SA only)		RESORCE[15]	Data for regorafenib presented in CS[1] Table 25
AE unit costs (SA only)		NHS Reference Costs 2019/20,[6] PSSRU[23] and assumptions	-

ERG - Evidence Review Group; PFS - progression-free survival; PAS - Patient Access Scheme; cPAS - comparator PAS; mg - milligram; RDI - relative dose intensity; SA - sensitivity analysis; BNF - British National Formulary; AE - adverse event; MAIC - matching-adjusted indirect comparison; OR - odds ratio; CS - company's submission

Table 15: Grade 3/4 AE frequency and unit costs (applied in sensitivity analysis 3 only)

AE	Unit cost	Frequency - cabozantinib†	Frequency - regorafenib*
PPES	£420.66	0.13	0.13
Hypertension	£638.81	0.55	0.13
Elevated aspartate aminotransferase	£0.00	0.11	0.05
Fatigue	£63.45	0.07	0.06
Diarrhoea	£629.69	0.12	0.02
Elevated bilirubin	£0.00	0.05	0.07
Expected cost per patient	-	£490.04	£155.86

AE - adverse event; PPES - palmar-plantar erythrodysaesthesia syndrome
** Calculated as the sum of Grade 3 and 4 treatment-emergent drug-related AEs in Bruix et al[15]*
† Calculated by applying the ORs from the company's MAICs to the regorafenib arm AE frequencies as baseline

4.2 Company's model results

The results of the company's base case analysis and sensitivity analyses are presented in Table 16. The company's base case analysis suggests that compared to regorafenib, cabozantinib is estimated to generate cost savings of [redacted] per patient. The estimated cost savings for cabozantinib are reduced slightly if patients spend less time on treatment and/or if the costs of managing AEs are included in the

analysis. The estimated cost savings are greater if patients spend longer on treatment, if RDI is included and/or if wastage costs are excluded from the model. The ERG notes that as these analyses do not include the cPAS discount for regorafenib, the results are not meaningful. The results of the company's model including the PAS discounted prices for cabozantinib and regorafenib are presented in a separate confidential appendix to this report.

Table 16: Results of company's cost comparison

Scenario	Cabozantinib	Regorafenib	Incremental
Base case	[redacted]	[redacted]	[redacted]
SA1 - Time on treatment – 20% ([redacted] months)	[redacted]	[redacted]	[redacted]
SA2 - Time on treatment + 20% ([redacted] months)	[redacted]	[redacted]	[redacted]
SA3 - Include arm-specific AE costs	[redacted]	[redacted]	[redacted]
SA4 - Include RDI	[redacted]	[redacted]	[redacted]
SA5 - Exclude wastage costs	[redacted]	[redacted]	[redacted]

SA - sensitivity analysis; AE - adverse event; RDI - relative dose intensity
** Regorafenib costs are unchanged from the base case due to patients spending 1-week off treatment at the end of each regorafenib treatment cycle (see clarification response,[19] question B5)*

4.3 ERG critique of the company's cost comparison model

4.3.1 Critical appraisal methods

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted cost comparison analysis. These included:

- Assessing whether the company's analysis is in line with NICE's guidance on undertaking cost comparison FTAs[24]
- Verifying the calculations used in the model, including double-programming the base case model and sensitivity analyses to check for errors
- Scrutinising the assumptions underpinning the cost comparison model and discussing these with clinical experts
- Checking the correspondence between the description of the model reported in the CS[1] and the company's executable model
- Where possible, checking of key parameter values used in the company's model against their original data sources.

As the company intends cabozantinib to be considered under NICE's FTA process, the focus of the ERG's critical appraisal was on the appropriateness of the cost comparison model and its underlying assumptions. The ERG's concerns around the submitted analysis are summarised briefly in Section 4.3.2. As discussed in Section 3.3, there is uncertainty around whether it is reasonable to assume clinical equivalence between cabozantinib versus regorafenib for PFS, OS and AEs. As such, the ERG's critique also includes some consideration of the likely direction of incremental costs and health outcomes if the assumption of equivalence does not hold.

4.3.2 ERG critical appraisal - results

The main items identified from the ERG's critique are summarised in Box 1.

Box 1: Summary of key items considered in the ERG's critical appraisal

(1) Adherence to NICE guidance on cost comparison FTAs
(2) Model verification
(3) Appropriateness of evidence sources
(4) Appropriateness of base case assumptions

(1) Adherence to NICE guidance on cost comparison FTAs

The company's cost comparison model includes a single comparator – regorafenib – which was appraised by NICE in TA514 and TA555.[5, 25] As discussed in Section B.1.1 of the CS,[1] the company's proposed positioning for cabozantinib is exactly the same as the current recommendation for regorafenib, that is, as an option for treating advanced unresectable HCC in adults who have had sorafenib, only if they have Child–Pugh grade A liver impairment and an ECOG PS of 0 or 1.[5] This is narrower than the wording of the marketing authorisation for cabozantinib for treating HCC,[11] although the ERG notes that all patients in CELESTIAL[7] had an ECOG PS <2 and only 1 patient had Child Pugh grade B disease. Given the company's intended positioning of cabozantinib, the ERG and its clinical advisors believe that the company's choice of comparator for the cost comparison is appropriate.

The final NICE scope[14] also includes BSC as a comparator. The CS[1] (Section B.1.1) comments that BSC *"is not a relevant comparator for a NICE FTA cost comparison for cabozantinib, as the comparator can only be technologies already recommended in published technology appraisal guidance and/or treatment guidelines for the same indication."*[1] The ERG agrees that BSC is not a relevant comparator for this appraisal and that a positive NICE recommendation for cabozantinib would only displace regorafenib.

The other aspects of the company's cost comparison analysis, including the time horizon adopted and the omission of discounting, are in line with NICE's guidance for companies submitting cost comparisons through the FTA process.[24]

(2) Model verification

The ERG double-programmed the company's cost comparison model. This included replicating the base case scenario and each of the five sensitivity analyses presented in Table 16. The ERG was able to generate the same results as those presented in the CS.[1] The ERG believes that the company's analyses are free from programming errors.

(3) Appropriateness of evidence sources used to inform model parameters

The ERG believes that the evidence sources used to inform the company's base case model (Table 14) are appropriate and that the values applied in the executable model are consistent with their original sources. The ERG also believes that the sources used to obtain these parameter values are appropriate. The ERG was unable to check whether the company's parametric survival modelling for PFS was implemented correctly as the underlying IPD were not provided.

The ERG notes that unit costs associated with managing AEs have been drawn from NHS Reference Costs 2019/20[6] and from the Personal Social Services Research Unit (PSSRU).[23] The most noticeable differences in AE frequencies between the drugs relate to hypertension and diarrhoea; other AE frequencies are similar between the groups (see Table 15). The unit cost for managing hypertension in the company's model is broadly similar to the value used in TA555[5] (cost comparison model unit cost = £629.69; TA555 model unit cost = £729.87), whilst diarrhoea was not included as an AE in the TA555 model. The ERG notes however that the general approach to modelling AEs differs between the appraisals – the cost comparison model assumes that AEs result in a once-only cost, whereas the TA555 model assumed an ongoing AE probability in every cycle at a lower overall rate.[26] As such, the approaches are therefore not fully comparable. However, neither the company's sensitivity analysis including differential AE costs (Table 16) nor the deterministic sensitivity analyses undertaken by the company in TA555 (see Stevenson *et al.*,[26] Figure 14) indicate that AE costs are a key model driver.

(4) Appropriateness of base case model assumptions

The ERG has some concerns regarding some of the base case model assumptions, in particular:

(a) The assumption of equivalent PFS and OS
(b) The assumption of equivalent AEs
(c) The assumption of equivalent resource use whilst on treatment
(d) The assumption of perfect (100%) RDI for both drugs.

These issues are discussed below.

(a) The assumption of equivalent PFS and OS

As discussed in Section 3.3, the company has undertaken a range of indirect comparisons using the Bucher approach and anchored and unanchored MAICs. All of these analyses suggest a statistically non-significant difference between cabozantinib and regorafenib for PFS and OS. The anchored MAICs, which reflect the preferred analyses of both the company and the ERG, indicate that the point estimate of the HR for PFS favours cabozantinib, whilst the point estimate of the HR for OS favours regorafenib. The ERG believes that there remains uncertainty around the relative treatment effect between cabozantinib and regorafenib, including the assumption of equivalence of the two regimens

and therefore, the results of the ITCs and the appropriateness of a cost comparison approach should be interpreted with caution.

(b) The assumption of equivalent AEs

The company's base case analysis excludes the costs of AEs. Given the use of a cost comparison approach, the analysis also assumes that there is no differential impact of toxicity on HRQoL between the treatment options. One of the ERG's clinical advisors commented that whilst the toxicity profile for regorafenib is both predictable and manageable, this is not the case for cabozantinib, which by comparison is considered to be less predictable and more toxic. This is likely to lead to increased costs and greater health losses for patients receiving cabozantinib compared with regorafenib. Differences in toxicity between the regimens are also apparent from the results of the company's MAICs, whereby the total sum of probabilities of the individual grade 3/4 AEs is 1.03 for cabozantinib and 0.46 for regorafenib, see Table 15). Whilst the company's sensitivity analyses include group-specific AE costs, the use of a cost comparison model precludes any consideration of associated health losses. Based on clinical advice received from clinical experts and the company's ITCs, the ERG believes that cabozantinib may result in QALY losses due to AEs, even if PFS and OS are broadly equivalent between the options. These effects cannot be fully captured in the company's cost comparison model.

The ERG also notes that the negative effects of toxicity may be reflected in the EQ-5D data from the CELESTIAL and RESORCE trials.[7, 15] The mean difference in change from baseline EQ-5D for regorafenib versus placebo in RESORCE was reported to be small and non-significant (mean difference in index score = -0.01; 95% CI -0.03 to 0.02, p=0.4695).[26] [redacted] Table 6 [redacted] Similarly, the ERG's clinical advisors highlighted the value that patients with advanced HCC place on maintaining HRQoL. One of the ERG's clinical advisors further commented that these toxicity effects are also evident from the data on dose reductions and Grade 3/4 AEs in the cabozantinib arm of CELESTIAL and the high proportion of Grade 3/4 AEs (62% of patients experienced a dose reduction and 68% of patients experienced Grade 3/4 AEs).

(c) The assumption of equivalent resource use whilst on treatment

The company's base case model assumes that all other resource use is equivalent for cabozantinib and regorafenib. The ERG's clinical advisors commented that owing to its comparatively worse toxicity profile, cabozantinib is expected to lead to additional costs of monthly face-to-face visits whilst patients are still on treatment, which would otherwise have been managed remotely and less frequently (2-monthly) for patients receiving regorafenib. These additional costs are not included in the company's base case or sensitivity analyses.

(d) The assumption of perfect RDI for both drugs

The company's base case analysis includes the costs of full pack dosing, based on the assumption that there are no efficiencies in minimising drug wastage in clinical practice (i.e., dose reductions, even if planned, do not lead to lower drug costs to the NHS). The CS[1] states that this approach reflects a conservative assumption and states that this assumption was used in TA555.[5] The ERG disagrees that this assumption was preferred in final guidance for TA555; the TA555 guidance document states that the company's analyses which assumed full pack dosing were *"unlikely to reflect clinical practice, because the dose reductions in the trial were planned, so it was more likely that wastage would be minimised in clinical practice"* (TA555 guidance, Section 3.15). As part of TA555, the company submitted evidence from pharmacists from two large tertiary centres in the UK supporting the use of pack-splitting to minimise wastage of sorafenib and other TKIs. The NICE Appraisal Committee concluded that *"although wastage could be minimised, the pharmacists' evidence provided by the company suggested that it could not be eliminated entirely"* Overall, the ERG believes that it may be more appropriate to include RDI, together with an assumed level of wastage which is consistent with previous appraisals in HCC.[10, 25]

4.5 Additional analyses undertaken by the company and the ERG

4.5.1 Additional analyses presented in the company's clarification response

During the clarification process, the ERG asked the company to fit parametric survival models to the OS data for the time-varying and constant HR anchored MAICs in the second-line HCC population and, if possible, to estimate incremental QALYs using these survival models together with the EQ-5D data from CELESTIAL[7] (see clarification response,[19] questions A22, B4 and B6). In their response, the company presented additional survival modelling, utility estimates based on CELESTIAL and a partitioned survival model which combines information on PFS, OS and utilities to estimate incremental QALYs for cabozantinib versus regorafenib. Incremental QALYs were presented across four scenarios:

1. Anchored MAIC, constant HR (Weibull HR). This model involved fitting parametric models for PFS and OS to data for each trial including treatment group as a covariate and applying the HR for regorafenib versus placebo to the weighted placebo arm of CELESTIAL. PFS and OS

were modelled using Weibull distributions. The company's clarification response indicates that this model reflects their base case scenario.

2. Anchored MAIC, constant HR (Cox PH). This model is the same as the company's base case, except that the HR from the Cox model used in the anchored MAIC was applied to the PFS and OS models for the cabozantinib group to estimate outcomes for the regorafenib group.
3. Anchored MAIC time-varying HR. This model applies time-varying HRs from the anchored MAICs. PFS and OS are both modelled using log-logistic models.
4. Unanchored MAIC. This model uses the unanchored MAIC, as described in Section B.3.10.3 of the CS.

For each of these four models, the company applied utility values for the progression-free and progressed disease states, based on a Tobit regression model fitted to the EQ-5D-5L data from CELESTIAL.[7] The same utility values were applied to each treatment group (utility value progression-free = [redacted]; utility value progressed disease = [redacted]). It should be noted that this approach implicitly assumes that cabozantinib is not associated with any further QALY losses due to toxicity compared to regorafenib. Incremental QALY estimates were presented using both the deterministic and probabilistic versions of the model.

The results of the company's partitioned survival models are summarised in Table 17. As expected, the company's anchored MAIC analyses, including their preferred base case, consistently indicate that cabozantinib is expected to result in an incremental QALY loss compared to regorafenib. In contrast, the unanchored MAIC indicates the reverse situation whereby cabozantinib results in an incremental QALY gain. The company's clarification response presents distributions of incremental QALYs from the probabilistic model and suggest that many probabilistic samples are close to zero, *"demonstrating no meaningful difference in QALYs between cabozantinib and regorafenib in a pure second line HCC population previously treated with sorafenib irrespective of tolerability."*[19] The ERG believes that the company's additional analyses are useful and that a good range of scenarios have been presented using appropriate methods. The ERG also agrees that the estimates of incremental QALYs are uncertain, but notes that if a full cost-utility model had been developed, the expected incremental QALYs would be negative and the resulting ICER would be in the North-West or South-West quadrants of the cost-effectiveness plane (depending on the discounted price of cabozantinib).

Table 17: Results of company's partitioned survival analysis

Scenario	Incremental QALYs gained - cabozantinib versus regorafenib	
	Deterministic model	Probabilistic model
1. Anchored MAIC, constant HR (Weibull HR base case)	[redacted]	[redacted]
2. Anchored MAIC, constant HR (Cox PH base case)	[redacted]	[redacted]
3. Anchored MAIC, time-varying HR	[redacted]	[redacted]
4. Unanchored MAIC	[redacted]	[redacted]

QALY - quality-adjusted life year; MAIC - matching-adjusted indirect comparison; HR - hazard ratio

4.5.2 Additional exploratory analysis undertaken by the ERG

In order to address some of the concerns raised in Section 4.4, the ERG undertook an additional exploratory analysis using the company's cost comparison model. This analysis is the same as the company's base case cost comparison, with the following amendments:

- RDI estimates are included, based on mean estimates reported from RESORCE and CELESTIAL[7, 15]
- AE management costs are included for both drugs
- Wastage costs are included based on two assumptions: (i) packs can be split to avoid inefficiencies in prescribing; (ii) on average, each patient will incur wastage associated with one quarter-pack of a pack of each drug (based on the earlier sorafenib HCC appraisal[10]).
- Monitoring costs are included for both drugs. For regorafenib, the analysis assumes that patients require one consultant-led non-face-to-face clinic visit every two months, whereas for cabozantinib, patients require one consultant-led face-to-face clinic visit every month. Unit costs were based on NHS Reference Costs 2019/20 (Consultant-led Medical Oncology, Service Code 370). The unit costs for non-face-to-face and face-to-face visits are £136.36 and £200.20, respectively.[6]

The results of the ERG's additional analysis are presented in Table 18. This analysis suggests slightly greater cost savings for cabozantinib, which are driven largely by the inclusion of RDI estimates in the analysis. In the absence of a full cost-utility model, the ERG is unable to undertake exploratory analyses under the assumption the cabozantinib and regorafenib are not equivalent in terms of PFS and OS. The ERG was also unable to undertake further analyses using the company's partitioned survival model described in the clarification response[19] (see Section 4.5.1) as the executable model was not provided.

Table 18: ERG's exploratory analyses using the company's cost comparison model

Scenario	Cabozantinib	Regorafenib	Incremental
Company's base case	[redacted]	[redacted]	[redacted]
ERG's preferred analysis under assumption of equivalence	[redacted]	[redacted]	[redacted]

4.6 ERG's view regarding whether outcomes and costs are likely to be similar for cabozantinib and regorafenib

Table 19 summarises the ERG's view regarding the direction of incremental health outcomes and costs, had a full cost-utility model been developed as part of a usual STA. Overall, the ERG believes that irrespective of whether it is reasonable to assume clinical equivalence in terms of PFS and OS, cabozantinib would likely be associated with fewer QALYs than regorafenib due to its comparatively worse toxicity. If relative treatment effects on clinical endpoints were based on the anchored MAICs, it is expected that cabozantinib would lead to a PFS gain and an OS loss; it is likely that the overall incremental health impact would be negative, as OS tends to have a greater impact on QALYs than PFS. If PFS is greater for cabozantinib than regorafenib, this would also likely lead to higher net drug acquisition costs, although this also depends on differences between the discounted prices of the two drugs. In the absence of a full cost-utility model, the magnitude of these expected QALY losses and cost differences remains unclear.

Table 19: Summary of ERG's view of the expected direction of incremental health outcomes and costs for cabozantinib versus regorafenib

Endpoint	ERG summary of evidence and comments
PFS	The company's Bucher ITCs and MAICs indicate non-significant differences in PFS. Point estimates of the HR are consistently in favour of cabozantinib. The ERG's clinical advisors commented that both drugs are likely to be similar in terms of PFS, but noted that the wide 95% CIs around the HRs means that there is uncertainty around the assumption of equivalence.
OS	The company's Bucher ITCs and MAICs indicate non-significant differences in OS. Point estimates of the HR are consistently in favour of regorafenib, except for the unanchored MAIC. As noted in the company's clarification response[19] (question B2), the proportions of patients receiving subsequent anticancer therapy in each trial was similar and is unlikely to confound OS results. The ERG's clinical advisors believe that both drugs are likely to be similar in terms of OS, but noted that the wide 95% CIs around the HRs means that there is uncertainty around the assumption of equivalence.
AE frequency	The company's MAICs indicate a greater overall incidence of Grade 3/4 AEs for cabozantinib than regorafenib (see Table 15). The ERG's clinical advisors commented that toxicity is worse for cabozantinib than regorafenib. One clinical advisor commented that this view reflects both the trial data and their own clinical experience with both drugs.
HRQoL	[redacted] Available EQ-5D data from RESORCE do not indicate a significant difference between regorafenib and placebo, which might suggest worse HRQoL for cabozantinib than regorafenib, although the EQ-5D questionnaire in RESORCE was completed on the first day of each treatment cycle, when a patient had not had treatment for a week, which may have affected patient responses. The ERG's clinical advisors commented that toxicity is worse for cabozantinib which likely means comparatively lower HRQoL.
Incremental QALYs	If a full cost-utility model had been developed using estimates of relative treatment effects from the anchored MAICs, regardless of toxicity effects, incremental QALYs for cabozantinib versus regorafenib would likely be negative, as OS tends to drive QALYs more than PFS. This can be seen in the company's partitioned survival analyses provided in their clarification response[19] (see Table 17). If PFS and OS were assumed to be equivalent, incremental QALYs for cabozantinib may still be negative due to toxicity effects. It is unclear whether the magnitude of these expected QALY losses would be sufficiently large to preclude cabozantinib from being considered under the FTA process.
Drug acquisition	In contrast to CELESTIAL and RESORCE, the ERG's clinical advisors commented that both cabozantinib and regorafenib would be given until disease progression. One advisor further commented that they would discontinue treatment if only if the patient had definite evidence of progression and if the patient was no longer benefiting from treatment. Time to treatment discontinuation (TTD) from either trial is therefore not a good proxy for ToT in clinical practice and the use of PFS is more appropriate. Differences in drug acquisition costs are dependent on the comparison of drug acquisition costs (including discounts) per period of time on treatment (see confidential appendix to this ERG report).
Drug administration	Not applicable - both drugs are administered orally.
Monitoring and health state costs	The company's cost comparison assumes no difference in costs of monitoring or visits. The ERG's clinical advisors commented that more frequent and less predictable AEs on cabozantinib would require patients to attend clinic in person, leading to increased costs.
AE costs	The ERG's clinical advisors believed that cabozantinib is more toxic than regorafenib. The costs of managing AEs are excluded from the company's base case analysis, but are included in sensitivity analysis. These costs are higher for cabozantinib than regorafenib and should be included in the analysis.
Incremental costs	Without a full cost-utility model, the incremental costs for cabozantinib versus regorafenib are not fully clear. If both drugs had the same acquisition cost per period of time on treatment, incremental costs for cabozantinib versus regorafenib would likely be slightly higher due to greater requirement to monitor and manage toxicity.

PFS - progression-free survival; OS - overall survival; AE - adverse event; HRQoL - health-related quality of life; QALY - quality-adjusted life year; HR - hazard ratio; ITC - indirect treatment comparison; MAIC - matching-adjusted indirect

comparison; ERG - Evidence Review Group; CSR - Clinical Study Report; TTD - time to treatment discontinuation; ToT - time on treatment

5 REFERENCES

1. Ipsen. Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]. Fast track appraisal: cost-comparison case. Company's submission Slough, UK; 2022.
2. Blachier M, Leleu H, Peck-Radosavljevic M, Valla DC, Roudot-Thoraval F. The burden of liver disease in Europe: a review of available epidemiological data. *Journal of Hepatology* 2013;58:593-608.
3. Cancer Research UK. Liver cancer incidence statistics. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/incidence#heading-Zero; 2022.
4. Cancer Research UK. Liver cancer mortality statistics. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/mortality#heading-Zero; 2022.
5. National Institute for Health and Care Excellence. Technology Appraisal Guidance 555 - Regorafenib for previously treated advanced hepatocellular carcinoma. London, UK; 2019.
6. National Health Service. NHS Reference Costs 2019/20. 2021.
7. Exelixis Inc. A Phase 3, randomized, double-blind, controlled study of cabozantinib (XL184) vs placebo in subjects with hepatocellular carcinoma who have received prior sorafenib (CELESTIAL Clinical Study Report). California, US; 2018.
8. National institute for Health and Care Excellence. Technology Appraisal Guidance 666 - Atezolizumab with bevacizumab for treating advanced or unresectable hepatocellular carcinoma. London, UK; 2020.
9. National institute for Health and Care Excellence. Technology Appraisal Guidance 551 - Lenvatinib for untreated advanced hepatocellular carcinoma London, UK; 2018.
10. National institute for Health and Care Excellence. Technology Appraisal Guidance 474 - Sorafenib for treating advanced hepatocellular carcinoma. London, UK; 2017.
11. European Medicines Agency. Cabometyx (cabozantinib) - Summary of Product Characteristics. London, UK; 2021.
12. Joint Formulary Committee. British National Formulary London, UK; 2022.
13. European Medicines Agency. Stivarga (regorafenib) - Summary of Product Characteristics London, UK; 2021.
14. National institute for Health and Care Excellence. Cabozantinib for previously treated advanced hepatocellular carcinoma (review of TA582). Final scope. London, UK; 2022.
15. Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, *et al.* Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 2017;389:56-66.
16. Ipsen. Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]. Fast track appraisal: cost-comparison case. Company's submission appendices. Slough, UK; 2022.
17. 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;379:54-63.
18. King J, Palmer DH, Johnson P, Ross P, Hubner RA, Sumpter K, *et al.* Sorafenib for the treatment of advanced hepatocellular cancer - a UK audit. *Clinical Oncology* 2017;29:256-62.
19. Ipsen. Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]. Company's response to clarification questions from the ERG. London, UK; 2022.

20. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* 1997;50:683-91.

21. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Medical Research Methodology* 2012;12.

22. Philippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Sheffield, UK; 2016.

23. Jones K, Burns A. Unit costs of health and social care. Kent, UK; 2021.

24. National institute for Health and Care Excellence. User guide for the cost comparison company evidence subnmission template. London, UK; 2017.

25. National Institute for Health and Care Excellence. NICE Technology Appraisal Guidance 514 - Regorafenib for previously treated advanced hepatocellular carcinoma. London, UK; 2017.

26. Stevenson M, Tappenden P, Carroll C, Ren S, Rawdin A, Wong R*, et al.* Regorafenib for previously treated unresectable hepatocellular carcinoma: A Single Technology Appraisal Sheffield, UK; 2017.

27. Ipsen. Ipsen Advanced Hepatocellular Carcinoma (HCC) Health Technology Assessment (HTA) Advisory Board. 28th June 2018. Meeting minutes. London, UK; 2018.

28. Ipsen. Ipsen NICE submission advisory board meeting. Hepatocellular carcinoma - cabometyx. London, UK; 2021.

Cabozantinib for previously treated advanced hepatocellular carcinoma (review of TA582). A Technology Appraisal

Addendum: Summary and critique of company's economic model

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 Rebecca Harvey, Statistician, Cabourn Statistics Ltd., France Katy Cooper, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK
Correspondence Author	Paul Tappenden, Professor of Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK
Date completed	30th August 2022 (post-FAC version)

1. Introduction

In May 2022, the National Institute for Health and Care Excellence (NICE) informed the company (Ipsen) that cabozantinib had failed the scrutiny stage of the NICE Fast Track Appraisal (FTA) process. Subsequently, it was agreed between NICE, the company and the Evidence Review Group (ERG) that a proportionate approach to the appraisal should subsequently be pursued. It was agreed that this would involve the company extending their existing partitioned survival model, which had previously been presented as part of the company's response to clarification questions from the ERG[1] (question B6), to estimate incremental quality-adjusted life years (QALYs) and incremental cost-effectiveness ratios (ICERs) for cabozantinib versus regorafenib. This model is discussed briefly in Section 4.5.1 of the ERG report.[2] In July 2022, the company provided an updated version of their submission to NICE[3] and a fully executable health economic model programmed in Microsoft Excel.®

This ERG addendum provides a summary and critique of the company's economic model and presents the results of additional exploratory analyses undertaken by the ERG. Several aspects of the updated company's submission (CS), including the indirect treatment comparisons (ITCs), remain unchanged from the original CS; hence, these are not discussed in detail in this addendum. The ERG's critique of these analyses can be found in Section 3.3 of the ERG report.[2]

All cost-effectiveness results presented in this addendum include the Patient Access Scheme (PAS) price for cabozantinib (discount=[redacted]) and the list price for regorafenib. The results of the economic analyses including the PAS discounts for both of these products is provided in a separate confidential appendix.

2. Description of company's model

2.1 Economic analysis scope

The scope of the company's economic model is summarised in Table 1. The model assesses the incremental cost-effectiveness of cabozantinib versus regorafenib in adult patients with advanced hepatocellular carcinoma (HCC) who have received prior sorafenib treatment and progressed following at least one prior systemic treatment. The analysis adopts a National Health Service (NHS) and Personal Social Services (PSS) perspective over a lifetime horizon. Health outcomes and costs are discounted at a rate of 3.5% per annum. In line with the ITCs summarised in the original CS[4] and the company's clarification response,[1] cost-effectiveness estimates for cabozantinib versus regorafenib are presented across three efficacy scenarios which reflect the anchored and unanchored matching-adjusted indirect comparisons (MAICs) based on time-to-event data from the CELESTIAL and RESORCE trials.[5, 6]

Table 1: Scope of company's additional economic analyses

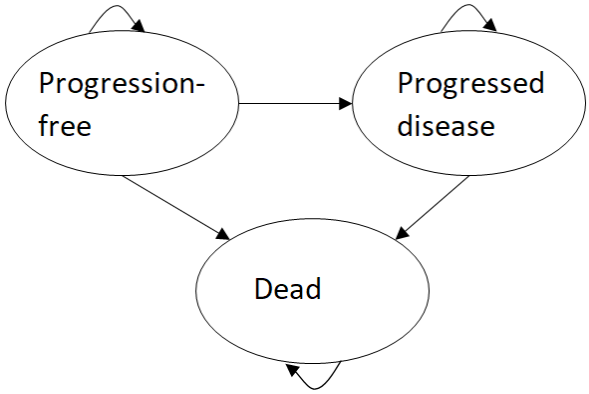
Population	Adult patients with advanced HCC who have received prior sorafenib treatment and progressed following at least one prior systemic treatment
Intervention	Cabozantinib 60mg QD
Comparator	Regorafenib 140mg QD for three weeks followed by one week off treatment
Outcome	Incremental cost per QALY gained
Time horizon	15 years (lifetime)
Perspective	NHS and PSS
Discounting	3.5% for health outcomes and costs
Efficacy scenarios considered	(1) Anchored MAIC, constant HRs (2) Anchored MAIC, time-varying HRs (3) Unanchored MAIC, independent models

HCC - hepatocellular carcinoma; mg - milligram; QD - once daily; QALY - quality-adjusted life year; NHS - National Health Service; PSS - Personal Social Services; MAIC - matching-adjusted indirect comparison; HR - hazard ratio

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 1).

Figure 1: Company's economic model structure



The model logic operates as follows. Patients enter the model in the progression-free state and receive treatment with either cabozantinib or regorafenib. At any time *t*, health state occupancy is determined by the cumulative probabilities of overall survival (OS) and progression-free survival (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. Both cabozantinib and regorafenib are assumed to be given until disease progression or death, whichever occurs first; hence, time to treatment discontinuation (TTD) is assumed to be equivalent to PFS. Patients in both treatment groups are assumed to also receive best supportive care (BSC) in every model cycle, regardless of whether they have progressed. Following disease progression, patients are assumed not to receive any further active anticancer therapy in either treatment group (i.e., patients receive BSC alone).

The cumulative probabilities of OS and PFS for patients receiving cabozantinib and regorafenib are estimated using parametric survival models fitted to the observed/MAIC-adjusted data from the CELESTIAL and RESORCE trials.[5, 6] The model applies a structural constraint whereby the cumulative probability of PFS cannot be higher than the cumulative probability of OS at any timepoint. No other structural constraints are included in the model.

Health-related quality of life (HRQoL) is assumed to be determined by the presence/absence of disease progression and the incidence of adverse events (AEs). The utility values applied in the progression-free and progressed disease states are based on a statistical model fitted to Euroqol 5-Dimensions 5-Level (EQ-5D-5L) data from CELESTIAL[5] (mapped to the 3-level [3L] version). The same utility values are applied in each treatment group. The model also applies AE-related QALY losses in every model cycle whilst the patient is progression-free. Utility values are not adjusted for increasing age.

The model includes costs associated with: (i) drug acquisition; (ii) disease management (health state costs); (iii) tests associated with disease progression; (iv) the management of AEs and (v) end-of-life care costs. Drug acquisition costs for cabozantinib and regorafenib are modelled as a function of the PFS distribution, the treatment schedule and daily dose, relative dose intensity (RDI) and the costs of each product (including the PAS price for cabozantinib and the list price for regorafenib). Costs associated with wastage are not included in the base case analyses. Health state costs are applied in each model cycle. Costs associated with AEs, disease progression and end-of-life care are applied once-only (in the first model cycle, at the point of progression and at the point of death, respectively).

The incremental health gains, costs and cost-effectiveness for cabozantinib versus regorafenib are estimated over a 15-year time horizon using a 28-day cycle duration. No economic subgroup analyses are presented in the CS.[3]

Cost-effectiveness results for cabozantinib versus regorafenib are presented across three efficacy scenarios which were previously presented in the original CS and clarification response:[1, 4]

1. Anchored MAIC, constant hazard ratios (HRs) for PFS and OS (Weibull models for both endpoints)
2. Anchored MAIC, time-varying HRs for PFS and OS (log-logistic models for both endpoints)
3. Unanchored MAIC, independently fitted PFS and OS models (generalised gamma models for PFS, log-logistic models for OS).

2.3 Key model assumptions

The company's model applies the following assumptions:

- The three efficacy scenarios presented in the updated CS[3] assume that cabozantinib is not clinically equivalent to regorafenib. The anchored MAICs (Efficacy Scenarios 1 and 2) apply

HRs which favour cabozantinib for PFS, but favour regorafenib for OS. The unanchored MAIC (Efficacy Scenario 3) applies independently fitted models which suggest that cabozantinib improves both PFS and OS compared with regorafenib.

- Patients are treated with regorafenib and cabozantinib until disease progression.
- All patients receive BSC in every model cycle.
- The model includes a constraint which ensures that the cumulative probability of PFS cannot be higher than the cumulative probability of OS. No other constraints are included.
- Excluding the impact of AEs, health state utility values for the progression-free and progressed disease states are assumed to be the same for both treatment groups.
- HRQoL impacts associated with AEs are applied in every model cycle, based on the frequency of AEs and the median treatment exposure time for cabozantinib and regorafenib. A single common disutility value is applied to all AEs.
- Costs associated with AEs are applied once only in the first cycle.
- The model assumes that disease management costs are lower for the progression-free state compared with the progressed disease state. The same disease management costs are applied to health states for each treatment group.
- The model also includes once-only costs of progression and death which are applied when patients leave the progression-free state and die, respectively.

2.4 Evidence used to inform the company's model parameters

Table 2 summarises the evidence sources used to inform the company's model parameters. The derivation of the model parameter values is discussed in the subsequent sections.

Table 2: Summary of evidence sources used to inform the company's model

Model parameter/group	Source
PFS and OS	MAICs of cabozantinib versus regorafenib using time-to-event data from CELESTIAL and RESORCE[5, 6]
TTD	Assumed to be equivalent to PFS
AE frequency	MAIC using data from CELESTIAL[5] and RESORCE[6] converted to per-cycle probability
Health state utility values	Multivariable Tobit regression with repeated measurements fitted to EQ-5D-5L data from CELESTIAL[5] (mapped to the 3L version using van Hout *et al.*[7])
AE disutility	
Amount of drug received	Dosing based on SmPCs for cabozantinib and regorafenib.[8, 9] RDI based on CELESTIAL and RESORCE.[5, 6] Wastage not included (assumes pack-splitting).
Other resource use	Based on survey of 30 HCC physicians (Li *et al.*[10])
End of life care costs	Coyle *et al.*[11]
Unit costs	BNF,[12] eMIT,[13] NHS Reference Costs 2019/20,[14] and the PSSRU[15]

PFS - progression-free survival; OS - overall survival; MAIC - matching-adjusted indirect comparison; TTD - time to treatment discontinuation; AE - adverse events; EQ-5D-5L - Euroqol 5-Dimensions 5-Levels; 3L - level; SmPC - summary of product characteristics; RDI - relative dose intensity; HCC - hepatocellular carcinoma; BNF - British National Formulary; NHS - National Health Service; PSSRU - Personal Social Services Research Unit; eMIT - electronic Market Information Tool

Time-to-event outcomes

The company's approach to modelling PFS and OS differs across each of the three efficacy scenarios:

- *Efficacy Scenario 1: Anchored MAIC, constant HR*. This approach involved fitting parametric models for PFS and OS to data for each trial including treatment group as a covariate and applying the HR for regorafenib versus placebo to the weighted placebo arm of CELESTIAL.[5] PFS and OS are modelled using Weibull distributions.
- *Efficacy Scenario 2: Anchored MAIC time-varying HR*. This scenario applies time-varying HRs from the anchored MAICs. PFS and OS are both modelled using log-logistic models.
- *Efficacy Scenario 3: Unanchored MAIC*. This scenario uses the unanchored MAIC, based on independently fitted models applied to the cabozantinib arm of CELESTIAL[5] and the regorafenib arm of RESORCE.[6]

These ITCs have been described and critiqued previously in Section 3.3 of the ERG report.[2] Kaplan-Meier plots, hazard plots and goodness of fit statistics (Akaike Information Criterion [AIC] and Bayesian Information Criterion [BIC] statistics) for the fitted parametric survival models are presented in the updated CS, the CS appendices and the clarification response.[1, 3, 16] The updated CS[3] states that parametric survival model selection was based on consideration of goodness-of-fit statistics, visual inspection and expert clinical input.[17, 18] A summary of the range of models considered, goodness-of-fit and clinical plausibility of the survival models fitted to the observed/MAIC-adjusted PFS and OS data is presented below.

Range of models assessed and goodness-of-fit

AIC and BIC statistics for the three efficacy scenarios can be found in CS Section B.3.3 (Tables 30, 31, 38 and 39) and CS Appendix L (Tables 51 and 52).[3, 16]

- *Efficacy Scenario 1 - Anchored MAIC, constant HR*
  - This analysis applies a constant HR to a baseline model; hence, the company only explored proportional hazards (PH) models within the analysis (the exponential, Weibull and Gompertz distributions).
  - The company selected the Weibull distribution for PFS and OS for both treatment groups. For both endpoints, the Weibull distribution is the best-fitting model in terms of both AIC and BIC for both treatment groups.
  - Based on visual inspection, the CS[3] comments that the PH assumption may not be appropriate and that modelled PFS and OS for the regorafenib group appear to be overestimated which biases against cabozantinib (see CS, Figures 32 and 35).

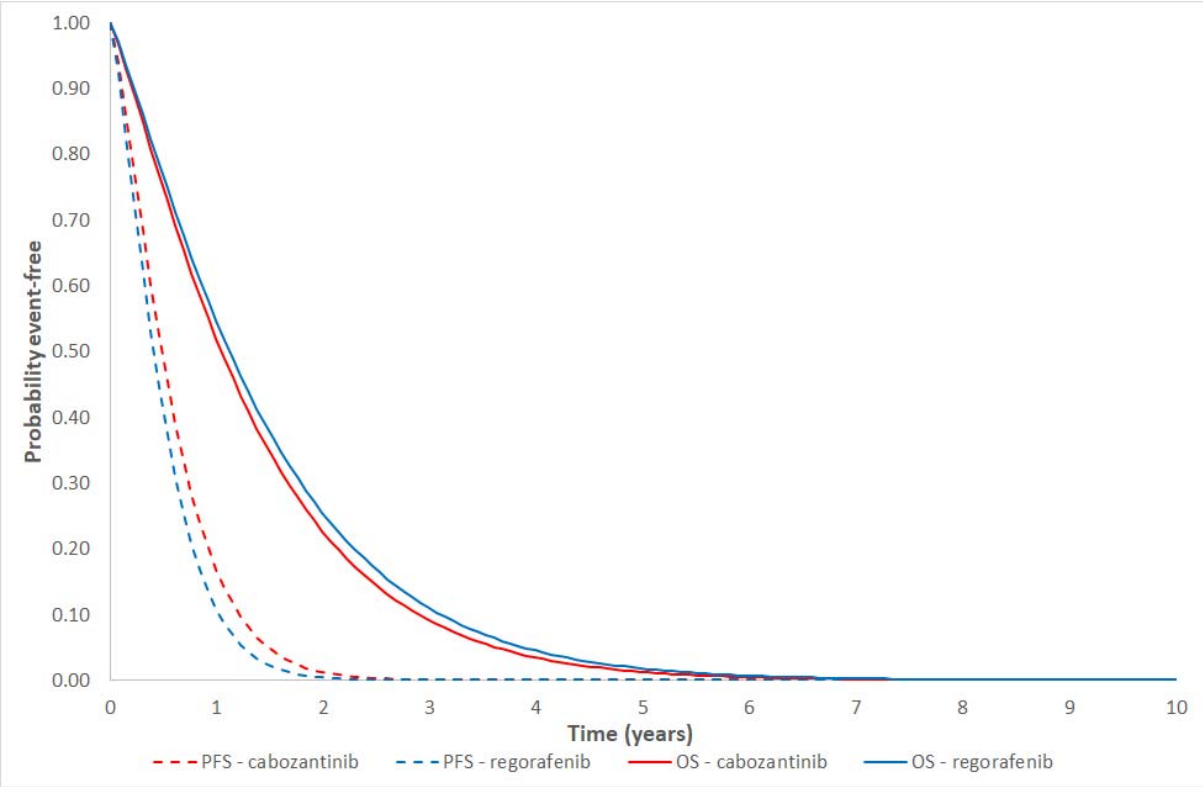
- *Efficacy Scenario 2 - Anchored MAIC, time-varying HR*
  - The company fitted six standard parametric survival models: the exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma distributions.
  - The company selected the log-logistic distribution for both PFS and OS for both treatment groups.
  - For PFS, the log-logistic distribution is the best-fitting model based on combined BIC and the second best-fitting model based on combined AIC.
  - For OS, the log-logistic distribution is the best-fitting model in terms of both AIC and BIC. The generalised gamma distribution has a similar AIC value, whilst the log-normal distribution has similar AIC and BIC values.
  - Based on visual inspection, the CS[3] comments that OS in the regorafenib group appears to be overestimated (see CS, Figure 33), but less so than in Efficacy Scenario 1 (anchored MAIC with constant HRs).

- *Efficacy Scenario 3 - Unanchored MAIC*
  - The company fitted six standard parametric survival models: the exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma distributions.
  - The company selected the generalised gamma model for PFS and the log-logistic model for OS. The same models are used in both treatment groups.
  - With respect to PFS, the generalised gamma distribution has the lowest AIC values. The log-logistic and log-normal models have lower BIC values for the cabozantinib and regorafenib arms, respectively. However, these differences are small.
  - With respect to OS, the log-logistic distribution is the best-fitting model for AIC and BIC in the cabozantinib arm, whereas the log-normal distribution is the best fitting model in the regorafenib arm.
  - The CS[3] does not comment on visual goodness of fit for this analysis; however, the ERG notes that the company's selected models appear to overestimate the tails of the distributions for the cabozantinib group, particularly for OS (see CS, Figures 25 and 26).

Summary of model-predicted PFS and OS

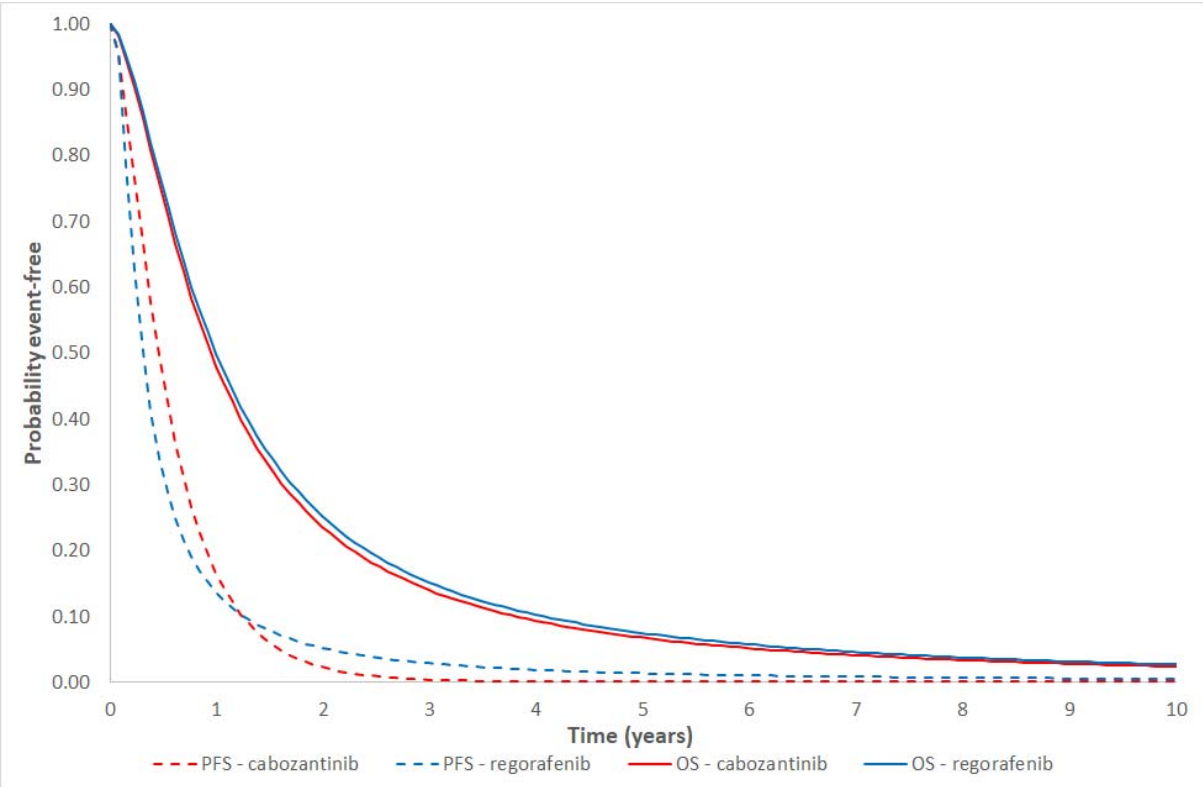
Model-predicted PFS and OS across the three efficacy scenarios are summarised in Figure 2, Figure 3 and Figure 4.

Figure 2: Modelled PFS and OS, Efficacy Scenario 1 – Anchored MAIC, constant HRs



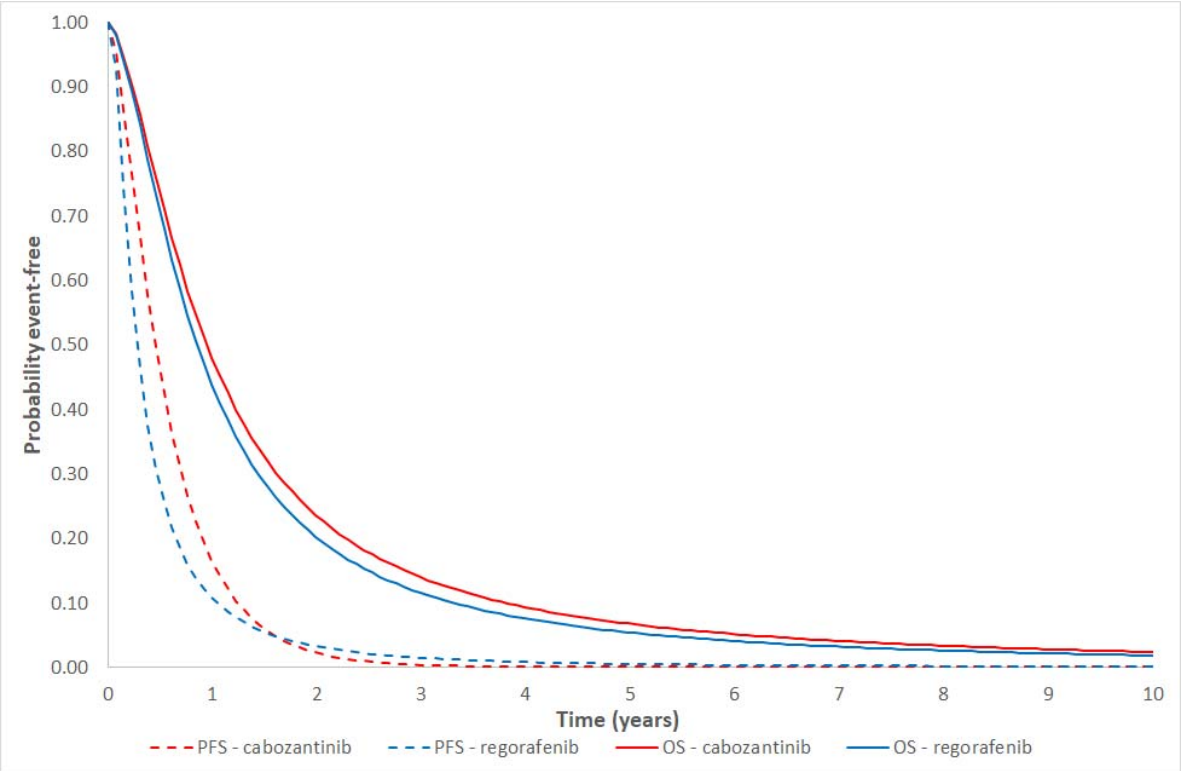
PFS - progression-free survival; OS - overall survival; MAIC- matching-adjusted indirect comparison; HR - hazard ratio

Figure 3: Modelled PFS and OS, Efficacy Scenario 2 – Anchored MAIC, time-varying HRs



PFS - progression-free survival; OS - overall survival; MAIC- matching-adjusted indirect comparison; HR - hazard ratio

Figure 4: Modelled PFS and OS, Efficacy Scenario 3 – Unanchored MAIC



PFS - progression-free survival; OS - overall survival; MAIC- matching-adjusted indirect comparison

Comparison of model predictions against external data

[redacted]

One of the ERG's clinical advisors commented that they would expect 4-year OS in patients treated with regorafenib to be less than 5% -

[redacted]

[redacted]. Efficacy Scenario 1 (anchored MAIC, constant HR) is broadly consistent with the ERG's clinical advisor's estimate, whilst the other two scenarios produce higher 4-year OS estimates of 8-10%. The limitations of each of the ITC methods should be considered when interpreting the results of each of the three efficacy scenarios (see Section 3.3 of the ERG report[2]). As discussed in the ERG report, the ERG considers the anchored MAIC analyses to provide the most robust estimates of relative treatment effects between cabozantinib and regorafenib; however, there are concerns regarding the comparability of the anchor arm (placebo plus BSC) across the CELESTIAL and RESORCE trials[5, 6] and the CS highlights potential bias regarding the overestimation of PFS and OS for the regorafenib group.

Table 3: Company's clinical experts' estimates of PFS and OS and company's model predictions

Efficacy scenario	Treatment group	PFS		OS
		2 years	4 year	4 years
Company's clinical experts	-	[redacted]	[redacted]	[redacted]
1. Anchored MAIC, constant HR	Cabozantinib	1%	0%	3%
	Regorafenib	0%	0%	5%
2. Anchored MAIC, time-varying HR	Cabozantinib	2%	0%	9%
	Regorafenib	5%	2%	10%
3. Unanchored MAIC	Cabozantinib	2%	0%	9%
	Regorafenib	3%	1%	8%

PFS - progression-free survival; OS - overall survival; MAIC - matching-adjusted indirect comparison; HR - hazard ratio

Frequency of AEs

The frequency of individual AEs for each treatment group are based on MAICs presented in Table 12 of the ERG report.[2] The model applies different approaches to estimate the impact of AEs on QALYs and costs:

- The model applies QALY losses associated with Grade 3/4 AEs in each model cycle in which the patient remains progression-free. The company estimated the per-cycle AE probability based on the overall proportion of patients experiencing any Grade 3/4 AE and the median treatment exposure time for cabozantinib and regorafenib in CELESTIAL and RESORCE.[5, 6]
- The model assumes that all costs associated with managing AEs are incurred in the first model cycle, based on the frequency of each individual AE and its respective cost.

Table 4: AE frequency and per-cycle probabilities applied in company's model

AE type	Cabozantinib (median treatment exposure time = [redacted] months)		Regorafenib (median treatment exposure time = 3.60 months)	
	Frequency	Cycle probability	Frequency	Cycle probability
PPES	0.13	0.03	0.13	0.04
Hypertension	0.55	0.16	0.13	0.04
Elevated AST	0.11	0.02	0.05	0.01
Fatigue	0.07	0.02	0.06	0.02
Diarrhoea	0.12	0.03	0.02	0.01

AE - adverse event; PPES - palmar-plantar erythrodysaesthesia; AST - aspartate aminotransferase

HRQoL

Health utility and disutility values were estimated using EQ-5D-5L data collected in CELESTIAL;[5] these data were mapped to 3-level (3L) version using the algorithm reported by van Hout *et al*.[7] The updated CS[3] states that the company explored several potential models to estimate utility values using the EQ-5D data, including ordinary least squares (OLS) regression, Tobit regression with repeated measurements and mixed models with repeated measurements. The final selected model is a multivariate Tobit regression model for repeated measurements; the CS states that this model was selected because it had a lower AIC value compared with the mixed model. This appears to be a similar statistical model to that described in the additional analyses presented in the company's clarification response[1] (question B4). The utility and disutility values applied in the company's economic model are summarised in Table 5.

Table 5: Utility and disutility values applied in company's model (adapted from CS, Table 45)

Health state	Mean value	SE
Utility - progression-free	[redacted]	[redacted]
Disutility - progressed disease	[redacted]	[redacted]
Disutility - AEs	[redacted]	[redacted]

SE - standard error; AE - adverse event

Resource use and costs

The model includes costs associated with: (i) drug acquisition; (ii) disease management (health state costs); (iii) tests associated with disease progression; (iv) the management of AEs and (v) end-of-life care costs. The costs applied in the company's economic model are summarised in Table 6. These are described in further detail in the subsequent sections.

Table 6: Summary of costs applied in the company's model

Cost type	Cabozantinib	Regorafenib
Drug acquisition costs (per 28 days, progression-free state only)	[redacted]	List price: £3,371.94
BSC costs (per 28 days, both health states)	£1.72	
Health state cost - progression-free (per 28 days)	£926.49	
Health state cost - progressed disease (per 28 days)	£1,362.60	
AEs (once-only)	£489.64	£155.86
Progression (once-only)	£627.87	
End of life care (once-only)	£5,818.34	

PAS - Patient Access Scheme; BSC - best supportive care; AE - adverse event

Drug acquisition costs

The drug acquisition costs applied in the model are shown in Table 7. Drug acquisition costs for cabozantinib and regorafenib are modelled as a function of the PFS distribution, the treatment schedule and daily dose, RDI and the costs of each product (including the PAS price for cabozantinib and the list price for regorafenib). Drug costs for cabozantinib and regorafenib were taken from the British National Formulary (BNF);[12] RDI was taken from the CELESTIAL and RESORCE trials.[5, 6] The base case model assumes that packs of cabozantinib and regorafenib can be split and that no tablets are wasted (every tablet prescribed is taken). As both drugs are taken orally, administration costs are not included in the model.

Table 7: Drug acquisition costs per 28 days

	Cabozantinib	**Regorafenib**
List price	£5,143.00	£3,744.00
Tablets per pack	30	84
RDI	0.61	0.90
PAS discount	[redacted]	Not included
Cost per 28-day cycle	[redacted]	**£3,371.94**

RDI - relative dose intensity; PAS - Patient Access Scheme

BSC costs

The model includes the costs of concomitant BSC including: cyclizine hydrochloride; dexamethasone; lactulose; metoclopramide; morphine sulphate; omeprazole; oramorph; paracetamol and spironolactone. All drugs were costed using prices from the Commercial Medicines Unit (CMU) Electronic Market Information Tool (eMIT).[13] Further details regarding the costs of individual BSC drugs can be found in Table 46 of the updated CS.[3] A total cost of £1.72 per 28-day cycle is applied to all patients in each model cycle.

Health state management costs

Health state costs applied in each model cycle are summarised in Table 8. These costs include hospitalisations, clinical consultations, laboratory tests, scans and radiotherapy. The proportions of patients and frequencies of each resource item per 28-day cycle were based on a survey of 30 physicians treating advanced HCC patients undertaken in 2018.[10] Unit costs were taken from the NHS Reference Costs 2019/20[14] and the Personal Social Services Research Unit (PSSRU).[15] Further details of the NHS Reference Cost service codes can be found in Table 48 of the updated CS.[3] The total health state costs per 28-day cycle were estimated to be £926.49 for patients who are progression-free and £1,362.60 for patients with progressed disease.

Table 8: Health state costs per 28-day cycle

Resource component	Unit cost	Progression-free				Progressed disease				Unit cost source
		No.	% patients	Duration (days)	Expected cost	No.	% patients	Duration (days)	Expected cost	
Hospitalisations										
General ward	£676.48	1.00	0.17	4.89	£566.32	1.00	5.36	0.27	£971.38	NHS Reference Costs 2019/20[14]
A&E admission	£205.09	0.70	0.20	1.00	£27.95	0.70	1.00	0.26	£37.72	
ICU	£270.61	1.00	0.03	3.50	£29.74	1.00	3.57	0.05	£48.69	
Medical staff visits										
Oncologist	£204.48	1.14	0.57	-	£131.96	0.96	0.63	-	£123.04	NHS Reference Costs 2019/20[14]
Hepatologist	£174.44	0.30	0.05	-	£2.62			-	£0.00	
Gastroenterologist	£154.41	0.44	0.22	-	£14.87	0.33	0.19	-	£9.75	
Clinical nurse specialist	£44.00	1.10	0.41	-	£19.76	1.00	0.42	-	£18.43	PSSRU[15]
Palliative care team	£44.00	0.33	0.30	-	£4.40	2.00	0.80	-	£70.40	
Macmillan nurse	£44.00	0.95	0.37	-	£15.52	1.22	0.42	-	£22.49	
General practitioner	£39.00	1.00	0.38	-	£14.97	0.96	0.42	-	£15.84	
Laboratory tests										
AFP test	£8.56	0.95	0.70	-	£5.65	0.91	0.66	-	£5.12	NHS Reference Costs 2019/20[14]
LFT	£8.56	1.09	0.78	-	£7.30	0.96	0.70	-	£5.75	
Biochemistry	£1.20	1.13	0.80	-	£1.08	1.00	0.71	-	£0.86	
Complete blood count	£2.27	1.13	0.79	-	£2.01	0.96	0.72	-	£1.56	
INR	£2.27	1.14	0.64	-	£1.64	1.05	0.62	-	£1.48	
Radiological tests										
CT scan	£123.71	0.88	0.51	-	£55.60	0.46	0.43	-	£24.25	NHS Reference Costs 2019/20[14]
MRI scan	£273.25	0.33	0.18	-	£16.17	0.06	0.12	-	£1.98	
Procedures										
Radiotherapy fraction	£739.30	0.26	0.05	-	£8.92	0.11	0.05	-	£3.86	NHS Reference Costs 2019/20[14]
Total health state cost	-	-	-		**£926.49**	-	-	-	**£1,362.60**	

A&E - accident and emergency; ICU - intensive care unit; AFP - alpha-fetoprotein; LFT - liver function test; INR - international normalised ratio; CT - computerised tomography; MRI - magnetic resonance imaging; NHS - National Health Service; PSSRU - Personal Social Services Research Unit

Costs associated with disease progression

The costs associated with disease progression are summarised in Table 9. These are assumed to include alpha-fetoprotein (AFP) tests, liver function tests (LFTs), computerised tomography (CT) scans and magnetic resonance imaging (MRI) scans. Resource usage was based on the physician survey[10] and unit costs were taken from NHS Reference Costs 2019/20.[14] These costs are applied once-only to the proportion of patients leaving the progression-free state in each model cycle.

Table 9: Disease progression costs (once-only)

Cost component	Unit cost	No.	Proportion patients	Expected cost
AFP test	£8.56	5.17	0.79	£34.93
LFT	£8.56	2	1.00	£17.13
CT scan	£123.71	7.4	0.61	£555.12
MRI scan	£273.25	0.35	0.22	£20.70
Total cost	-	-	-	**£627.87**

AFP - alpha-fetoprotein; LFT - liver function test; CT - computerised tomography; MRI - magnetic resonance imaging

AE management costs

Costs associated with AEs are summarised in Table 10. The frequency of AEs was estimated using the company's MAICs (see ERG report,[2] Table 12). Unit costs were based on NHS Reference Costs 2019/20.[14] These costs are applied once-only in the first model cycle.

Table 10: AE costs (once-only)

AE	Unit cost	Frequency cabozantinib	Frequency regorafenib
PPES	£420.66	0.13	0.13
Hypertension	£638.81	0.55	0.13
Elevated AST	£0.00	0.11	0.05
Fatigue	£63.45	0.07	0.06
Diarrhoea	£629.69	0.12	0.02
Elevated bilirubin	£0.00	0.05	0.07
Expected cost	-	**£489.64**	**£155.86**

AE - adverse event; palmar-plantar erythrodysaesthesia; AST - aspartate aminotransferase

End of life care costs

The model includes a cost associated with end-of-life care of £5,818.34. This value was taken from Coyle *et al.*[11] and was uplifted to current values using inflation indices from the PSSRU.[15]

2.5 Model evaluation methods

The updated CS[3] presents base case cost-effectiveness results for each of the three efficacy scenarios using both the deterministic and probabilistic versions of the model. The probabilistic ICER is based on 1,000 Monte Carlo samples. The results of the probabilistic sensitivity analysis (PSA) for all efficacy scenarios are also presented using a cost-effectiveness plane and cost-effectiveness acceptability curves

(CEACs). The updated CS[3] presents the results of deterministic sensitivity analyses (DSAs) using tornado plots. The CS also reports the results of a range of deterministic scenario analyses exploring alternative assumptions regarding: the time horizon; treatment duration; the exclusion of RDI; discount rates; the use of list prices for both drugs; alternative parametric survival models; the use of Bucher ITCs rather than MAICs; the inclusion of wastage costs and alternative health state utility values.

2.6 Company's model results

Table 11 presents the central estimates of cost-effectiveness generated using the company's model across the three efficacy scenarios. All results include the PAS for cabozantinib and the list price for regorafenib. The results of the probabilistic analyses indicate that using the anchored MAICs, cabozantinib is expected to generate fewer QALYs and incur lower costs than regorafenib; the probabilistic ICERs are large and are in the South-West quadrant. The unanchored MAIC suggests that cabozantinib is expected to generate additional QALYs and cost-savings; hence, cabozantinib dominates regorafenib. The results generated using the deterministic version of the model for Efficacy Scenarios 1 and 3 are generally similar to those obtained from the probabilistic model; the probabilistic results for Efficacy Scenario 2 (MAIC with time-varying HR) suggest greater expected QALY losses and cost savings compared with the deterministic model.

Table 11: Summary of company's base case cost-effectiveness results

Option	LYGs*	QALYs	Costs	Inc. LYGs*	Inc. QALYs	Inc. costs	ICER
1. CEA, anchored MAIC, constant HRs (probabilistic)†							
Cabozantinib	1.43	[redacted]	[redacted]	-0.09	[redacted]	[redacted]	**£295,334 (SWQ)**
Regorafenib	1.53	1.05	£55,001	-	-	-	-
2. CEA, anchored MAIC, time-varying HRs (probabilistic) †							
Cabozantinib	1.81	[redacted]	[redacted]	-0.14	[redacted]	[redacted]	**£224,469 (SWQ)**
Regorafenib	1.95	1.27	£60,303	-	-	-	-
3. CEA, unanchored MAIC (probabilistic)†							
Cabozantinib	1.82	[redacted]	[redacted]	0.21	[redacted]	[redacted]	**Dominating**
Regorafenib	1.62	1.07	£55,409	-	-	-	-
1. CEA, anchored MAIC, constant HRs (deterministic)							
Cabozantinib	1.42	[redacted]	[redacted]	-0.10	[redacted]	[redacted]	**£290,383 (SWQ)**
Regorafenib	1.52	1.04	£55,669	-	-	-	-
2. CEA, anchored MAIC, time-varying HRs (deterministic)							
Cabozantinib	1.81	[redacted]	[redacted]	-0.10	[redacted]	[redacted]	**£300,170 (SWQ)**
Regorafenib	1.90	1.25	£60,496	-	-	-	-
3. CEA, unanchored MAIC (deterministic)							
Cabozantinib	1.81	[redacted]	[redacted]	0.19	[redacted]	[redacted]	**Dominating**
Regorafenib	1.62	1.07	£56,058	-	-	-	-

LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; CEA - cost-effectiveness analysis; MAIC - matching-adjusted indirect comparison; HR - hazard ratio; SWQ – South-West quadrant
** Undiscounted*
† Based on a re-run of the probabilistic version of the model by the ERG, using 10,000 Monte Carlo simulations

Summary of other uncertainty analyses presented in the updated CS

The company's tornado plots for each efficacy scenario are presented in Figures 42, 43 and 44 of the updated CS.[3] These plots present the incremental net monetary benefit (NMB) for cabozantinib versus regorafenib assuming a willingness-to-pay (WTP) threshold of £30,000 per QALY gained. For brevity, these are not reproduced here. The company's plots consistently indicate that cabozantinib generates more NMB than regorafenib across all analyses, with the daily cost of regorafenib being the most influential model driver across all three efficacy scenarios.

The company's cost-effectiveness planes and CEACs for all three efficacy scenarios are presented in Figures 45 and 46 of the updated CS, respectively.[3] Assuming a WTP threshold of £30,000 per QALY gained, the probability that cabozantinib generates more net benefit than regorafenib is estimated to be approximately 0.94 or higher.

The results of the company's scenario analyses are summarised in Table 61 of the updated CS.[3] For brevity, these are not reproduced here. The economic conclusions suggested by these analyses are similar to those of the company's base case analyses (see Table 11), with the following exceptions:

- Using the list price for both cabozantinib and regorafenib results in substantially less favourable ICERs for cabozantinib (Efficacy Scenario 1: £25,227 per QALY gained [SWQ]; Efficacy Scenario 2: Dominated; Efficacy Scenario 3: £30,255 per QALY gained).
- The Bucher ITC results suggest that cabozantinib generates fewer QALYs and saves costs compared with regorafenib, leading to a South-West quadrant ICER of £162,411 per QALY gained.

These analyses indicate that the relative effectiveness of cabozantinib versus regorafenib and the prices of these products are key model drivers.

3. Critical appraisal by the ERG

3.1 Critical appraisal methods

The ERG 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.[19, 20]
- Scrutiny of the company's model by the ERG.
- 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 updated CS and the company's executable model.
- Replication of the base case results and PSA using the company's executable model.
- Where possible, checking of key parameter values used in the company's model against their original data sources.
- Clinical expert input to assess the plausibility of the model predictions.

3.2 Model verification

The ERG double-programmed the deterministic version of the company's model in order to check its implementation across all three efficacy scenarios. The results of the ERG's double-programmed model are very similar results to those generated using the company's model. During the process of rebuilding the company's model, the ERG identified several errors and other minor issues; these are described in Section 3.5.

Table 12: Comparison of results generated using the company's model and the ERG's double-programmed model, deterministic

Model outcome (incremental)	Company's model	ERG's double-programmed model
Efficacy scenario 1. Anchored MAIC, constant HRs		
Inc. LYGs	-0.10	-0.10
Inc. QALYs	[redacted]	[redacted]
Inc. costs	[redacted]	[redacted]
ICER	**£290,383 (SWQ)**	**£290,382 (SWQ)**
Efficacy scenario 2. Anchored MAIC, time-varying HRs		
Inc. LYGs	-0.10	-0.10
Inc. QALYs	[redacted]	[redacted]
Inc. costs	[redacted]	[redacted]
ICER	**£300,170 (SWQ)**	**£300,168 (SWQ)**
Efficacy scenario 3. Unanchored MAIC		
Inc. LYGs	0.19	0.19
Inc. QALYs	[redacted]	[redacted]
Inc. costs	[redacted]	[redacted]
ICER	**Dominating**	**Dominating**

ERG - Evidence Review Group; Inc. - incremental; LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; MAIC - matching-adjusted life year; HR - hazard ratio; SWQ - South-West quadrant

3.3 Adherence to the NICE Reference Case

Table 13 summarises the extent to which the company's model adheres to the NICE Reference Case.[21] The ERG has no major concerns and considers that the company's model is in line with the Reference Case.

Table 13: Adherence to the NICE Reference Case

Element of HTA	Reference Case	ERG comments
Defining the decision problem	The scope developed by NICE	The model compares cabozantinib against regorafenib in adult patients with advanced HCC who have had sorafenib. The final NICE scope[22] includes BSC as a comparator but this is not included in the economic model. As discussed in the ERG report,[2] the ERG agrees that BSC is not a relevant comparator for the population in whom regorafenib would otherwise be used.
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 gains accrued by patients.
Perspective on costs	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 adopts a 15-year time horizon. Across all three efficacy scenarios, virtually all patients (>98.5%) in the model have died by the final model cycle.
Synthesis of evidence on health effects	Based on systematic review	Modelled health outcomes have been estimated using ITCs comparing cabozantinib versus regorafenib using data from CELESTIAL and RESORCE.[5, 6] These trials were identified by the company's clinical effectiveness SLR.[3]
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 and a disutility value associated with AEs have been estimated using EQ-5D-5L data collected in CELESTIAL[5] (mapped to the 3L version using the algorithm reported by Van Hout *et al.*[7]).
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	No additional QALY weighting is applied.
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	The model includes costs borne by the NHS and PSS, valued using NHS Reference Costs and other standard costing sources.
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 annum.

HTA - health technology assessment; ERG - Evidence Review Group; NICE - National Institute for Health and Care Excellence; HCC - hepatocellular carcinoma; PSS - Personal Social Services; BSC - best supportive care; ITC - indirect treatment comparison; QALY - quality-adjusted life year; EQ-5D - Euroqol 5-Dimensions; 5L - 5-level; HRQoL - health-related quality of life; SLR - systematic literature review; AE - adverse event

3.4 Correspondence between model parameter values and evidence sources

Where possible, the ERG checked the parameter values used in the company's model against their original sources. The company's parametric survival models, HRs and HRQoL model were derived using individual patient data (IPD) which were not made available to the ERG; as such, the ERG cannot verify that these values have been estimated appropriately.

The ERG notes the following potential concerns regarding the other model parameters:

- The ERG was unable to find the number of patients attending A&E departments from the physician survey poster reported by Li *et al.*[10]
- The model worksheet "Cost inputs" suggests that the number of scans and tests incurred on disease progression were derived from the physician survey. However, these values are not reported by Li *et al.*[10] As such, the source of these values is unclear.
- The ERG was unable to identify or derive the company's unit cost estimates for hospitalisations from the NHS Reference Costs.[14]

The ERG believes that these issues are likely to be minor. The ERG was able to identify or derive all other cost and resource estimates used in the company's model.

3.5 Other issues identified from the ERG's critical appraisal

Other issues identified from the ERG's critical appraisal are summarised in Box 1. These issues are discussed below.

Box 1: Issues identified by the ERG's critical appraisal

(1) Model errors and other problems
(2) Issues relating to model parameter values
(3) Assumption of equivalent health state costs for cabozantinib and regorafenib
(4) Exclusion of wastage costs
(5) Discrepancy between probabilistic and deterministic results for Efficacy Scenario 2

(1) Model errors and other problems

The ERG identified five issues in the implementation of the company's model:

(a) The company's half-cycle correction is applied incorrectly as the first cycle is counted 1.5 times, rather than 0.5 times. This overestimates costs and health outcomes in both treatment groups.
(b) Costs associated with progression and end-of-life care are calculated based on the half-cycle corrected model trace. The ERG believes that it would be more appropriate to use the uncorrected trace for these costs.

(c) The physician survey poster (Li *et al.*[10]) reports resource use estimates per month, but the company's model applies these estimates in each 28-day model cycle. These costs should have been adjusted to reflect the 28-day cycle length (i.e., multiplied by 28/30.44).
(d) The model does not include a general population constraint.
(e) The model does not include age-adjustment of utility values or a cap to ensure that the modelled utility values for people with HCC remain lower than those for the general population.

These issues are addressed as part of the ERG's additional exploratory analyses (see Section 4).

(2) Issues relating to model parameter values

The ERG believes that the evidence sources used to inform the model parameters are generally appropriate.

The ERG does not have any major concerns regarding the company's survival analysis or model selection process, and the ERG broadly agrees with the final selected models included in each of the three efficacy scenarios. The three efficacy scenarios generate model predictions of PFS and OS which are broadly consistent with the views of clinical experts consulted by the company (see Table 3). Efficacy Scenario 1 appears to be most consistent with the ERG's clinical advisor's expectations of 4-year OS. The company has noted that OS appears to be overestimated in the regorafenib group in Efficacy Scenarios 1 and 2, whilst the ERG notes that OS appears to be overestimated for the cabozantinib group in Efficacy Scenario 3.

With respect to the HRQoL parameters, the ERG does not have any major methodological concerns regarding the company's analysis of the EQ-5D data from CELESTIAL,[5] but notes that the estimated disutility value associated with disease progression appears low (disutility = [redacted]). One of the ERG's clinical advisors commented that they would expect HRQoL to deteriorate more rapidly in patients with disease progression than in patients who are receiving an effective treatment - this deterioration does not appear to be fully reflected in the EQ-5D estimates used in the model. As such, the utility value for the progressed disease state (utility value = [redacted]) may not fully reflect the average level of HRQoL experienced by patients with advanced HCC who have failed two TKIs over their entire remaining lifetime. The ERG notes however that the post-progression utility values applied in the models used to inform NICE TA474[23] and TA514[24] also applied relatively high post-progression utility values based on analyses of EQ-5D data collected in the SHARP and RESORCE trials[6, 25] (utility values of 0.71 and 0.76, respectively). The ERG's exploratory analyses include a sensitivity analysis using a larger disutility value to explore its impact on the cost-effectiveness results (see Section 4).

The ERG also notes that the Coyle *et al.* study,[11] which used to inform the costs of end-of-life care, is more than 20 years old and that more recent sources are available (e.g., Round *et al.*[26]). However,

because virtually all patients in the model incur this cost, and most patients have a short survival time, this parameter has very little impact on the model results.

(3) Assumption of equivalent health state costs for cabozantinib and regorafenib

The company's model assumes that disease management costs in the progression-free health state are equivalent for cabozantinib and regorafenib. The ERG's clinical advisors commented that owing to its comparatively worse toxicity profile, cabozantinib is expected to lead to additional costs of monthly face-to-face visits whilst patients are still on treatment, which would otherwise have been managed remotely and less frequently (2-monthly) for patients receiving regorafenib. These additional costs are not included in the company's base case or sensitivity analyses. The ERG's exploratory analyses include additional monitoring costs for cabozantinib (see Section 4).

(4) Exclusion of wastage costs

The company's base case analyses assume that packs of treatment can be split and that every tablet prescribed is taken; hence, no wastage costs are included. This assumption particularly advantages the cabozantinib group because the mean RDI is much lower than that for regorafenib (0.61 vs 0.90). The ERG notes that some patients will incur wastage because they progress or die before completing a pack of treatment. The ERG believes that it would be more appropriate to include a level of drug wastage which is consistent with previous appraisals in HCC.[23, 24] These costs have been included in the ERG's exploratory analyses (see Section 4).

(5) Discrepancy between probabilistic and deterministic results for Efficacy Scenario 2

As shown in Table 11, the results of the probabilistic and deterministic results for the MAIC with time-varying HRs are noticeably different, with the former suggesting a comparatively greater loss in survival and QALYs than the latter. The ERG scrutinised the company's PSA sampling sub-routine and believes that this apparent discrepancy is due to uncertainty around the sampled survival model parameters rather than being the consequence of an error. Whilst the PSA results presented in the CS^3 are based on 1,000 Monte Carlo samples, all probabilistic results reported in this addendum use 10,000 samples.

4. Additional exploratory analyses undertaken by the ERG

4.1 ERG exploratory analysis - methods

The ERG undertook six sets of exploratory analyses (EAs) using the deterministic version of the company's model. These analyses are described below.

EA1: Correction of errors

This analysis includes the correction of three errors in the company's model:

(a) The half-cycle correction calculations were amended to count the first model cycle 0.5 times rather than 1.5 times.

(b) The calculations relating to the costs of progression and death were amended to use the uncorrected model trace.
(c) The health state cost calculations were amended to reflect a 28-day cycle duration.

These corrections were applied in all subsequent exploratory analyses.

EA2: Include general population mortality constraint

A general population mortality constraint was applied to the OS models to ensure that the risk of death with the disease in each cycle cannot be lower than the risk of all-cause death in the age- and sex-matched general population. This was done using a weighted survival model based on general population life tables for England,[27] together with information on the median age and proportion of female patients in the CELESTIAL trial (age=64 years; proportion female=0.18).[5]

EA3: Inclusion of age-adjusted utilities

Utility values were adjusted for increasing age based on a multiplicative approach using EQ-5D-3L estimates reported by Hernandez Alava *et al.*[28]

EA4: Inclusion of additional monitoring costs for cabozantinib

The health state cost calculations for the cabozantinib group were amended to include the cost of 0.5 additional oncologist visits per month (0.46 visits per 28-day model cycle).

EA5: Inclusion of wastage costs

The model was amended to include the costs of 7 days' worth of treatment in both treatment groups (adjusted for RDI). This was implemented using existing functionality contained in the company's model.

EA6: ERG-preferred model

The ERG's preferred model includes EA1-5. Results of this exploratory analysis are presented using both the deterministic and probabilistic versions of the model.

Additional sensitivity analyses

The ERG undertook four sets of additional sensitivity analyses using the ERG's preferred model (EA6).

ASA1: Alternative PFS models

The model was re-run using all alternative PFS models.

ASA2: Alternative OS models

The model was re-run using all alternative OS models.

ASA3: Post-progression utility value doubled

The disutility value associated with disease progression was doubled.

4.2 ERG exploratory analysis – results

ERG's preferred model results

The results of the ERG's preferred analyses for each of the three efficacy scenarios are presented in Table 14. The ERG's preferred model using the anchored MAICs suggests that compared with regorafenib, cabozantinib generates fewer QALYs and saves costs, leading to a high South-West quadrant ICERs of £254,307 and £202,316 saved per QALY lost for Efficacy Scenarios 1 and 2, respectively. The ERG's preferred model using the unanchored MAIC (Efficacy Scenario 3) suggests that cabozantinib generates additional QALYs and reduces costs, thereby dominating regorafenib.

Table 14: ERG preferred model results

Analysis	Incremental - cabozantinib versus regorafenib			
	Inc. LYGs	Inc. QALYs	Inc. costs	ICER
Efficacy scenario 1 – Anchored MAIC, constant HR				
Company's base case (deterministic)	-0.10	[redacted]	[redacted]	**£290,383 (SWQ)**
EA1 - Correction of errors	-0.10	[redacted]	[redacted]	£252,357 (SWQ)
EA2: General population mortality constraint	-0.10	[redacted]	[redacted]	£252,357 (SWQ)
EA3: Age-adjusted utilities	-0.10	[redacted]	[redacted]	£254,180 (SWQ)
EA4: Additional monitoring visit cost	-0.10	[redacted]	[redacted]	£241,519 (SWQ)
EA5: Wastage included	-0.10	[redacted]	[redacted]	£260,606 (SWQ)
EA6a: ERG-preferred model (deterministic)	-0.10	[redacted]	[redacted]	£251,572 (SWQ)
EA6b: ERG-preferred model (probabilistic)	-0.09	[redacted]	[redacted]	**£254,307 (SWQ)**
Efficacy scenario 2 – Anchored MAIC, time-varying HR				
Company's base case (deterministic)	-0.10	[redacted]	[redacted]	**£300,170 (SWQ)**
EA1 - Correction of errors	-0.10	[redacted]	[redacted]	£257,547 (SWQ)
EA2: General population mortality constraint	-0.10	[redacted]	[redacted]	£257,547 (SWQ)
EA3: Age-adjusted utilities	-0.10	[redacted]	[redacted]	£261,597 (SWQ)
EA4: Additional monitoring visit cost	-0.10	[redacted]	[redacted]	£243,674 (SWQ)
EA5: Wastage included	-0.10	[redacted]	[redacted]	£266,626 (SWQ)
EA6a: ERG-preferred model (deterministic)	-0.10	[redacted]	[redacted]	£256,727 (SWQ)
EA6b: ERG-preferred model (probabilistic)	-0.14	[redacted]	[redacted]	**£202,316 (SWQ)**
Efficacy scenario 3 – Unanchored MAIC				
Company's base case (deterministic)	0.19	[redacted]	[redacted]	**Dominating**
EA1 - Correction of errors	0.19	[redacted]	[redacted]	Dominating
EA2: General population mortality constraint	0.19	[redacted]	[redacted]	Dominating
EA3: Age-adjusted utilities	0.19	[redacted]	[redacted]	Dominating
EA4: Additional monitoring visit cost	0.19	[redacted]	[redacted]	Dominating
EA5: Wastage included	0.19	[redacted]	[redacted]	Dominating
EA6a: ERG-preferred model (deterministic)	0.19	[redacted]	[redacted]	Dominating
EA6b: ERG-preferred model (probabilistic)	0.21	[redacted]	[redacted]	**Dominating**

LYG - life year gained; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; EA - exploratory analysis; ERG - Evidence Review Group; MAIC - matching-adjusted indirect comparison; HR - hazard ratio

ERG's additional sensitivity analysis results

The results of the ERG's additional sensitivity analyses are summarised in Table 15. The economic conclusions remain consistent across all additional sensitivity analyses.

Table 15: ERG additional sensitivity analysis results

Analysis	**ICER – cabozantinib versus regorafenib**		
	1. Anchored MAIC, constant HR	**2. Anchored MAIC, time-varying HR**	**3. Unanchored MAIC**
ERG preferred model (deterministic)	**£251,572 (SWQ)**	**£256,727 (SWQ)**	**Dominating**
ASA1 – PFS = exponential	Not modifiable. Model uses Weibull distributions for PFS and OS.	£304,858 (SWQ)	Dominating
ASA1 – PFS = Weibull		£276,427 (SWQ)	Dominating
ASA1 – PFS = Gompertz		£282,716 (SWQ)	Dominating
ASA1 – PFS = log-normal		£243,518 (SWQ)	Dominating
ASA1 – PFS = log-logistic		£256,727 (SWQ)	Dominating
ASA1 – PFS = generalised gamma		£330,385 (SWQ)	Dominating
ASA2 – OS = exponential		£496,592 (SWQ)	Dominating
ASA2 – OS = Weibull		£297,850 (SWQ)	Dominating
ASA2 – OS = Gompertz		£64,981 (SWQ)	Dominating
ASA2 – OS = log-normal		£226,129 (SWQ)	Dominating
ASA2 – OS = log-logistic		£256,727 (SWQ)	Dominating
ASA2 – OS = generalised gamma		£132,798 (SWQ)	Dominating
ASA3 – progression disutility doubled	£271,009 (SWQ)	£292,878 (SWQ)	Dominating

ICER - incremental cost-effectiveness ratio; MAIC - matching-adjusted indirect comparison; HR - hazard ratio; ERG - Evidence Review Group; ASA - additional sensitivity analysis; PFS - progression-free survival; OS - overall survival; RDI - relative dose intensity; SWQ - South-West quadrant

5. End of life

The updated CS[3] states that cabozantinib does not meet NICE's End of Life criteria. The ERG agrees with the company's view.

6. References

1. Ipsen. Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]. Company's response to clarification questions from the ERG. Ipsen: Sough, UK; 2022.
2. Tappenden P, Harvey R, Cooper K, Wong R. Cabozantinib for previously treated advanced hepatocellular carcinoma (review of TA582). A Technology Appraisal. Evidence Review Group report. University of Sheffield: Sheffield, UK; 2022.
3. Ipsen. Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]. Company's updated submission. Ipsen: Slough, UK; 2022.
4. Ipsen. Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]. Fast track appraisal: cost-comparison case. Company's submission. Ipsen: Slough, UK; 2022.
5. Exelixis Inc. A Phase 3, randomized, double-blind, controlled study of cabozantinib (XL184) vs placebo in subjects with hepatocellular carcinoma who have received prior sorafenib (CELESTIAL Clinical Study Report). Ipsen: California, US; 2018.
6. Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, *et al.* Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 2017;389:56-66.
7. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, *et al.* Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value in Health* 2012;15:708-15.
8. European Medicines Agency. Cabometyx (cabozantinib) - Summary of Product Characteristics. EMA: London, UK; 2021.
9. European Medicines Agency. Stivarga (regorafenib) - Summary of Product Characteristics EMA: London, UK; 2021.
10. Li YMS, Mollon P, Guerra I. Healthcare resource use among patients with advanced hepatocellular carcinoma (aHCC) in the UK. HCC-UK Annual Conference. 2020.
11. Coyle D, Small N, Ashworth A, Hennessy S, Jenkins-Clarke S, Mannion R, *et al.* Costs of palliative care in the community, in hospitals and in hospices in the UK. *Critical Reviews in Oncology/Hematology* 1999;32:71-85.
12. Joint Formulary Committee. British National Formulary. BMJ Group and Pharmaceutical Press: London, UK; 2022.
13. Commercial Medicines Unit. Drugs and pharmaceutical electronic market information tool (eMIT). CMU. 2022.
14. National Health Service. NHS Reference Costs 2019/20. NHS. 2021.
15. Jones K, Burns A. Unit costs of health and social care. PSSRU: Kent, UK; 2021.
16. Ipsen. Cabozantinib for previously treated advanced hepatocellular carcinoma [ID3917]. Company's updated submission appendices. Ipsen: Slough, UK; 2022.
17. Ipsen. Ipsen Advanced Hepatocellular Carcinoma (HCC) Health Technology Assessment (HTA) Advisory Board. 28th June 2018. Meeting minutes. Ipsen: Slough, UK; 2018.
18. Ipsen. Ipsen NICE submission advisory board meeting. Hepatocellular carcinoma - cabometyx. Ipsen: Slough, UK; 2021.
19. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th edn. New York: Oxford University Press; 2015.
20. 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.

21. National institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. NICE: London, UK; 2013.
22. National institute for Health and Care Excellence. Cabozantinib for previously treated advanced hepatocellular carcinoma (review of TA582). Final scope. NICE: London, UK; 2022.
23. National institute for Health and Care Excellence. Technology Appraisal Guidance 474 - Sorafenib for treating advanced hepatocellular carcinoma. NICE: London, UK; 2017.
24. National Institute for Health and Care Excellence. NICE Technology Appraisal Guidance 514 - Regorafenib for previously treated advanced hepatocellular carcinoma. NICE: London, UK; 2017.
25. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, *et al.* Sorafenib in advanced hepatocellular carcinoma. *New England Journal of Medicine* 2008;359:378-90.
26. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. *Palliative Medicine* 2015;29:899-907.
27. Office for National Statistics. National life tables: England 2018-2020; 2021.
28. Hernandez Alava M, Pudney S, Wailoo A. Estimating EQ-5D by age and sex for the UK. DSU: Sheffield, UK; 2022.