Regorafenib for previously treated unresectable hepatocellular carcinoma: A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK Paul Tappenden, Reader in Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK Christopher Carroll, Reader in Systematic Review and Evidence Synthesis, ScHARR, University of Sheffield, Sheffield, UK Shijie Ren, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Andrew Rawdin, Research Assistant, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK Suzanne Darby, Consultant Medical Oncologist, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield Michael Heneghan, Consultant Hepatologist and Professor of Hepatology, King's College Hospital, London
Correspondence Author	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK

Date completed	14th September 2017

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 16/108/02.

Declared competing interests of the authors

None of the authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Hazel Squires, ScHARR, for providing comments on the draft report and Gill Rooney, 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 HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Stevenson M., Tappenden P, Carroll C, Ren S, Rawdin A, Wong R, Darby S, Heneghan M. Regorafenib for previously treated unresectable hepatocellular carcinoma: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2017.

Contributions of authors

Ruth Wong critiqued the company's search strategy. Chris Carroll summarised and critiqued the clinical effectiveness data reported within the company's submission. Paul Tappenden, Andrew Rawdin and Matt Stevenson critiqued the health economic analysis submitted by the company and performed the ERG exploratory analyses. Dr Darby and Professor Heneghan provided clinical advice to the ERG. All authors were involved in drafting and commenting on the final report.

Standard copyright statement on the front page of the ERG/AG report:

Copyright belongs to ScHARR

Copyright is retained by Bayer for Figures 1,4,5,6,7,10,11,13,14,15,16 and 17 and Tables 2,4,8,9,10 and 23.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

A&E	Accident and emergency
AACR	American Association for Cancer Research
AASLD	American Association for the Study of Liver Diseases
AE	Adverse events
AFP	Alpha-fetoprotein
AFT	Accelerated Failure Time
AIC	Akaike Information Criterion
ALT	Alanine aminotransferase
ANC	Absolute neutrophil count
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
AUC	Area under the curve
BIC	Bayesian Information Criterion
BP	Blood pressure
BSC	Best Supportive Care
CDF	Cancer Drugs Fund
CEAC	Cost-effectiveness acceptability curve
CI	Confidence interval
CR	Complete response
CRD	Centre for Reviews and Dissemination
CRF	Case report form
CS	Company's submission
CSCO	Chinese Society of Clinical Oncology
CSR	Clinical Study Report
CT	Computerised tomography
DCO	Data cut-off
DCR	Disease control rate
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
EASL-EORTC	European Association for Study of the Liver / European Organisation for the Research and Treatment of Cancer
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EMBASE	Excerpta Medica dataBASE
EQ-5D	EuroQol – 5 Dimensions
ERG	Evidence Review Group
ESDO	European Society of Digestive Oncology
ESMO	European Society for Medical Oncology
FACT-G	Functional Assessment of Cancer Therapy – General
FACT-Hep	Functional Assessment of Cancer Therapy – Hepatobiliary
FAD	Final Appraisal Determination
FCE	Finished consultant episode
HCC	Hepatocellular carcinoma
HR	Hazard ratio
HRG	Healthcare resource group
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio

ILCA	International Liver Cancer Association
IPD	Individual patient-level data
IQR	Interquartile range
ITT	Intention-to-treat
IVRS	Interactive voice response system
JSMO	Japanese Society of Medical Oncology
LSM	Least-Squares Mean
LYG	Life year gained
MedDRA	Medical Dictionary for Regulatory Activities
MEDLINE	Medical Literature Analysis and Retrieval System Online
MID	Minimally important difference
mRECIST	Modified Response Evaluation Criteria In Solid Tumors
MRI	Magnetic resonance imaging
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NLM	National Library of Medicine
NYHA	New York Heart Association
OLS	Ordinary least squares
ORR	Objective response rate
OS	Overall survival
PAS	Patient Access Scheme
PD	Progressed disease
PFS	Progression-free survival
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RESORCE	Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment
SAE	Serious adverse event
SD	Standard deviation
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
TACE	Transarterial chemoembolisation
TESAE	Treatment-emergent severe AEs
TKI	Tyrosine kinase inhibitor
TOI	Trial Outcome Index
TTP	Time to progression
ULN	Upper limit of normal
VAS	Visual analogue scale
WTP	Willingness-to-pay

1 SUMMARY

1.1 Critique of the decision problem in the company's submission

The company's submission (CS) adequately describes the decision problem. The CS assesses the clinical and cost-effectiveness of regorafenib (Stivarga®), within its licensed indication for the treatment of adult patients with hepatocellular carcinoma (HCC) who have been previously treated with sorafenib. The positioning of regorafenib within the treatment pathway was appropriately reserved for patients who have received sorafenib treatment, and the comparator of best supportive care (BSC) was appropriate. Evidence relating to all outcomes listed in the final scope produced by the National Institute for Health and Care Excellence (NICE) was included within the CS.

1.2 Summary of clinical effectiveness evidence submitted by the company

The CS identified a single, relevant study: the RESORCE trial. This was an international, placebo-controlled Phase III trial which evaluated the efficacy and safety of regorafenib 160mg per day in adult patients with HCC who have previously progressed on sorafenib. In terms of the primary outcome, the RESORCE study found that patients on regorafenib had increased survival: the median overall survival (OS) was reported to be 10.6 months (95% CI 9.1-12.1 months) in patients randomised to regorafenib compared with 7.8 months (95% CI 6.3-8.8 months) in patients randomised to placebo. The estimated hazard ratio (HR) for OS for regorafenib compared with placebo was 0.63 (95% confidence interval [CI] 0.50-0.79, one-sided p=0.000020).

The CS also reported the secondary and tertiary outcomes of the RESORCE trial. Median progression-free survival (PFS), as measured by modified response evaluation criteria in solid tumors (mRECIST), was significantly better for regorafenib (3.1 months, 95% CI 2.8–4.2 months) than for placebo (1.5 months, 95% CI 1.4–1.6 months): HR, 0.46, 95% CI 0.37-0.56; $p<0.0001$. The median time to progression (TTP) as measured by mRECIST was also significantly better for regorafenib (3.2 months, 95% CI 2.9–4.2 months) than for placebo (1.5 months, 95% CI 1.4–1.6 months): HR, 0.46, 95% CI 0.36-0.55; $p<0.0001$. The objective response rate (ORR), which aggregates complete response (CR) and partial response (PR) according to mRECIST, was also significantly higher in the regorafenib group than the placebo group (11% compared with 4%; $p=0.0047$). Similar findings were reported across all outcomes when using the slightly different RECIST 1.1 criteria. Subgroup analyses demonstrated consistent benefit for patients treated with regorafenib, although an additional pre-specified analysis found that those who develop a new extrahepatic lesion when they progressed on sorafenib had a considerably worse survival rate compared with those who did not. The RESORCE trial also found that health-related quality of life (HRQoL) was similar between the groups, but was consistently worse for regorafenib than placebo across different measures. These differences were found to be statistically significant in the case of the Functional Assessment of Cancer Therapy – Hepatobiliary (FACT-Hep)

total and the Trial Outcome Index (TOI), but did not reach clinical significance according to pre-specified thresholds.

Adverse events (AEs) were frequent: 100% of regorafenib patients receiving the study drug experienced at least one AE (compared with 93% on placebo), and 93% of regorafenib patients experienced treatment-emergent drug-related AEs compared with 52% of placebo patients. The principal AEs were: hand foot skin reaction (53% in the regorafenib arm compared with 8% in the placebo arm); diarrhoea (41% vs 15%); fatigue (40% vs 32%); hypertension (41% vs 6%); and anorexia (31% vs 15%). AEs of Grade 3 or higher were reported for 80% of patients in the regorafenib group compared with 59% in the placebo group. More regorafenib patients than placebo patients also experienced Grade 3 (46% compared with 16%) and Grade 4 (4% compared with 1%) drug-related AEs. The incidence of haemorrhage events of ≥ Grade 3 was higher in the placebo group (8%) than the regorafenib group (6%), but the incidence of drug-related haemorrhage events of ≥ Grade 3 was higher in the regorafenib group (1.6%) than the placebo group (0%). According to the CS, the incidence of drug-related severe adverse events (SAEs) was relatively low in both groups, but was higher in regorafenib-treated patients compared with those receiving placebo (10% vs. 3%).

Sixty-eight percent of regorafenib patients had dose interruptions or reductions due to AEs compared with 31% of placebo patients, and dose interruptions or reductions due to drug-related AEs occurred in 54% of regorafenib patients and 10% of placebo patients. According to the CS, dose reductions (not including interruptions) due to AEs occurred in [redacted] of the patients in the regorafenib group and [redacted] of the patients in the placebo group. The AE profile of regorafenib in the RESORCE trial is generally similar to trials of regorafenib undertaken in patients with colorectal cancer. Deaths assessed as related to the study drug were reported for seven (2%) regorafenib patients and two (1%) placebo patients. There are no relevant ongoing studies of regorafenib.

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

The company's systematic review was generally well conducted. However, some processes could have been reported better and some relevant abstracts and additional analyses relating to the pivotal RESORCE trial should have been identified and included in the CS. This additional literature is cited, where appropriate, throughout this ERG report. The included relevant study, the RESORCE trial, is a high quality randomised controlled trial (RCT), with a low risk of selection, performance, detection, attrition and reporting bias.

The principal issue with the evidence relates to the generalisability of the trial population to the population of patients seen in clinical practice in the UK. The RESORCE trial only included meaningful data on patients who were not found to be intolerant to sorafenib, who had an Eastern Cooperation

Oncology Group (ECOG) Performance Status (PS) of 0 or 1, and who were categorised as Child-Pugh class A, whilst the marketing authorisation for regorafenib covers all adult patients with hepatocellular carcinoma (HCC) who have been previously treated with sorafenib (even if they are found to be intolerant to sorafenib, or are ECOG PS 2 or Child-Pugh class B). A recent audit of sorafenib use in the UK found that sorafenib is used in patients who are ECOG PS 2 and Child-Pugh class B (21% and 16% of the audit population, respectively). These patients have a poorer prognosis than patients enrolled in the RESORCE trial. There is therefore a lack of clinical data on the efficacy and safety of regorafenib in these groups - this issue is acknowledged in the CS. This is important because the sorafenib audit found that ECOG PS $\geq$2 was an independent predictor of mortality and OS was substantially worse in patients who were Child-Pugh class B (4.6 months) compared with those who were Child-Pugh class A (9.5 months). Pre-specified subgroup analyses conducted using data from RESORCE also found that patients who were PS 0 and Child-Pugh A5 experienced better efficacy than those who were PS 1 and Child-Pugh A6. The sorafenib audit also reported that liver dysfunction was much more common as an AE in Child-Pugh class B patients (40%) compared with Child-Pugh class A patients (18%), as was deterioration in PS (47% vs 32%). It is therefore possible that patients treated in UK clinical practice may experience less efficacy and more AEs than patients enrolled in RESORCE. The lack of relevant data and its implications are acknowledged in the Summary of Product Characteristics (SmPC) for regorafenib, which recognises the potential adverse impact of regorafenib on hepatic function in patients who are Child-Pugh class B and the need to monitor all AEs carefully in this group. There is therefore substantial uncertainty concerning the benefits of regorafenib in patients who do not satisfy the inclusion criteria of the RESORCE trial.

1.4 Summary of cost effectiveness submitted evidence by the company

The company submitted a health model constructed in Microsoft Excel®. The model adopts a partitioned survival approach based on three health states: (1) progression-free; (2) progressed disease, and (3) dead. The time horizon was approximately 15 years with 28-day cycles. The clinical parameters of the model were informed by analyses of time-to-event data (PFS, OS and time on treatment) collected within the RESORCE trial. Resource use and unit costs were drawn from the RESORCE trial and other sources, including a survey of three leading clinical experts. Based on the deterministic version of the company's original submitted model, the incremental cost effectiveness ratio (ICER) for regorafenib versus BSC was estimated to be £33,437 per quality-adjusted life year (QALY) gained. Following the clarification process, two further versions of the model were submitted by the company. The company's revised base case analysis, which includes longer-term data corresponding to the 23rd January 2017 data cut-off (DCO), dependent log normal OS curves and a truncated log logistic time to treatment discontinuation function, produces a deterministic ICER for regorafenib versus BSC of £36,050 per QALY gained.

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

The ERG critically appraised the company's economic analysis and double-programmed the deterministic version of the company's original submitted model. The ERG's critical appraisal identified several issues relating to the company's economic analysis and the evidence used to inform it. The most pertinent of these include: (i) the inappropriate use of a hazard ratio (HR) to model relative treatment effects on OS; (ii) limited consideration of the clinical plausibility of the extrapolated OS curves; (iii) concerns regarding the modelling of time to treatment discontinuation; (iii) the inclusion of potentially unrealistic cost savings due to dose reductions and treatment interruptions; (iv) the use of the 2015 survey of three experts to inform health state resource use (and the exclusion of the earlier survey used to inform the recent sorafenib appraisal); (v) concerns regarding the appropriateness of several unit cost estimates; (vi) the questionable reliability of the post-progression utility estimate and (vii) the inadequate representation of parameter uncertainty.

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

1.6.1 Strengths

The systematic review presented within the CS has been undertaken to a good standard. The ERG considers the RESORCE study to be a high quality RCT.

With the exception of the approach adopted to model time spent receiving regorafenib, the ERG considers the general model structure adopted by the company to be appropriate.

1.6.2 Weaknesses and areas of uncertainty

There is an absence of trial evidence on some patient groups who would be eligible to receive regorafenib: adults with HCC who are sorafenib intolerant or who are Child-Pugh class B or who have ECOG PS 2.

The rationale for some of the assumptions used within the company's model were unclear or contentious. Many of these assumptions were favourable to regorafenib; when alternative more appropriate parameter values are used, the ICER for regorafenib increases substantially.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG performed seven sets of exploratory analyses to explore the impact of alternative assumptions on the ICER. These analyses involve: (1) the correction of unequivocal model errors and use of alternative unit costs; (2) the inclusion of a more appropriate general ward day bed cost; (3) the use of full pack dosing which does not include cost savings due to reduced dosing; (4) the removal of half-cycle correction for drug acquisition costs; (5) the use of combined 2007 and 2015 survey costs (as preferred by the Cancer Drugs Fund [CDF] Appraisal Committee within the recent appraisal of

sorafenib for HCC); (6) the use of independent Weibull functions to model OS, and (7) the use of a fully extrapolated log logistic time to treatment discontinuation curve (with full pack dosing). These exploratory analyses were then combined to form the ERG's preferred base case (analysis 8).

The results of the ERG's exploratory analyses are presented in Table 1. The ERG's preferred base case deterministic ICER for regorafenib versus BSC is £81,081 per QALY gained. The ERG notes that the ICER would increase slightly if a greater disutility for progression disease is assumed. The ERG also notes that where a reduction in dose is planned and the lower dose is to be maintained over the long-term, the ERG's assumption of indefinite full pack dosing for all patients will lead to an overestimation of the ICER for regorafenib. Additional sensitivity analyses undertaken by the ERG indicate that even under the highly optimistic assumption that all patients have indefinite dose reductions to [redacted] from the start of treatment, the ICER for regorafenib versus BSC remains above [redacted] per QALY gained.

Table 1: Exploratory analyses undertaken by the ERG and the ERG-preferred base case

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Company's base case (revised base case model, deterministic)					
Regorafenib	1.073	[redacted]	0.406	£14,625	**£36,050**
BSC	0.668	[redacted]	-	-	-
Exploratory analysis 1: Correction of unequivocal model errors and use of alternative unit costs					
Regorafenib	1.048	[redacted]	0.368	£12,659	**£34,406**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 2: Inclusion of more appropriate general ward bed day cost**					
Regorafenib	1.048	[redacted]	0.368	£12,647	**£34,376**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 3: Use of full pack dosing**					
Regorafenib	1.048	[redacted]	0.368	£15,508	**£42,151**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 4: Removal of half-cycle correction for drug acquisition costs**					
Regorafenib	1.048	[redacted]	0.368	£13,332	**£36,235**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 5: Use of combined 2007 and 2015 survey costs**					
Regorafenib	1.048	[redacted]	0.368	£20,297	**£55,166**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 6: Use of independent Weibull functions to model OS**					
Regorafenib	0.896	[redacted]	0.265	£10,242	**£38,683**
BSC	0.632	[redacted]	-	-	-
*Exploratory analysis 7: Use of a fully extrapolated log logistic time to discontinuation curve (patients on treatment at 29th February 2016 censored, with full pack dosing)**					
Regorafenib	1.048	[redacted]	0.368	£21,751	**£59,120**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 8: ERG's preferred base case (including all individual amendments)**					
Regorafenib	0.896	[redacted]	0.265	£21,468	**£81,081**
BSC	0.632	[redacted]	-	-	-

2 BACKGROUND

2.1 Critique of company's description of the underlying health problem

The company's submission[1] (CS) provides an adequate description of hepatocellular carcinoma (HCC) which includes stating that: it the 17th most common cancer in the UK; that it affects more men than women, and that incidence of the disease increases with age.[2] HCC is stated to be "*often diagnosed at a late stage of the disease when patients present with symptoms including fatigue, jaundice, pruritus, encephalopathy, weight loss, ascites, abdominal pain / distension and the presence of a mass*." (CS,[1] page 19).

Figure 1 of the CS provides the classification of HCC using the joint European Association for Study of the Liver / European Organisation for the Research and Treatment of Cancer (EASL-EORTC) guidelines.[3] The company also present a table representing the staging of HCC using the Barcelona Clinic Liver Cancer (BCLC) classification and how this relates to Eastern Cooperative Oncology Group (ECOG) performance status (PS) and Child-Pugh class. These data are reproduced in Figure 1 and Table 2, respectively.

Figure 1: Classification of HCC (from EASL-EORTC Clinical Practice Guidelines)

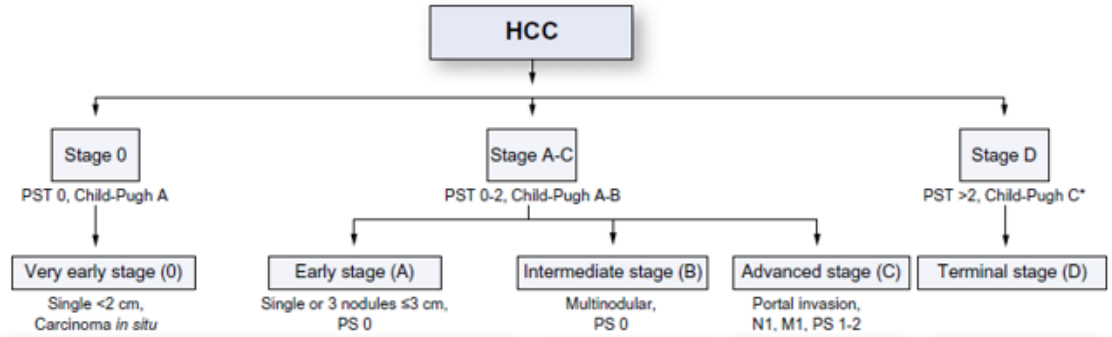


Table 2: Staging of HCC (using the BCLC classification)

BCLC Stage	Tumour status	ECOG performance status	Liver Function (Child-Pugh)
0 (Very early HCC)	Single tumour < 2cm in diameter without vascular invasion / satellites	0	Well preserved liver function Child-Pugh A
A (Early HCC)	Single tumours >2 cm or 3 nodules <3 cm of diameter	0	Child–Pugh A or B
B (Intermediate HCC)	Multinodular asymptomatic tumours without an invasive pattern	0	Child-Pugh A-C
C (Advanced HCC)	Symptomatic tumours; macrovascular invasion (either segmental or portal invasion) or extrahepatic spread (lymph node involvement or metastases)	1–2	Child-Pugh A-C
D (End stage HCC)	Tumours leading to a very poor performance Status which reflects a severe tumour-related disability	3-4	Child-Pugh C

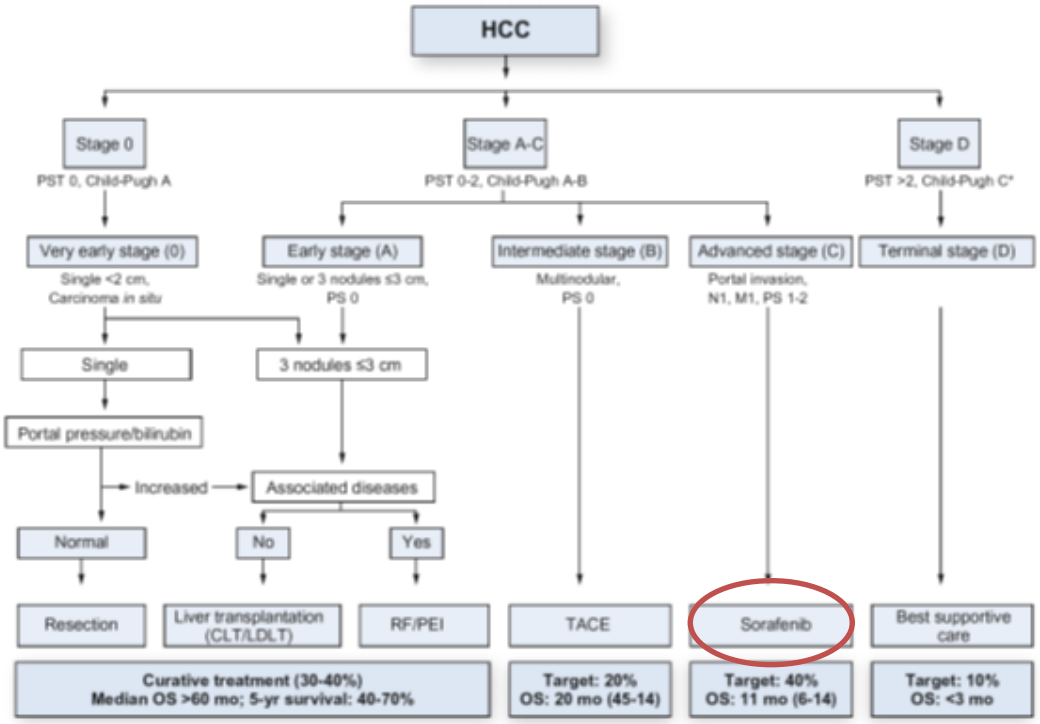
Whilst the CS summarises expected Child-Pugh grade in terms of the BCLC stage, it does not detail how the Child-Pugh classification was estimated. A full description may be of value to the Appraisal Committee as clinical advice received by the Evidence Review Group (ERG) suggests that there is little difference between a person with a Child-Pugh score of 6 (which is classified as an A) and a person with a Child-Pugh of 7 (which is classified as a B). The Child-Pugh score is generated from five clinical measures of liver disease: (i) total bilirubin; (ii) serum albumin; (iii) prothrombin time; (iv) ascites, and (v) hepatic encephalopathy. Each measure is scored between one and three (with a score of three indicating greater severity), thereby resulting in an overall score between five and fifteen. Scores of 5 or 6 are classified as Child-Pugh A, scores of 7, 8 or 9 are classified as Child-Pugh B and scores of ten and over are classified as Child-Pugh C.[4] Further details are provided in Section 4.2.1.

2.2 Critique of company's overview of current service provision

The CS[1] provides a satisfactory overview of current service provision. The CS states that as UK-specific guidelines are dated, these have been largely superseded by the EASL-EORTC guidelines for the treatment of advanced HCC. Within these guidelines, choice of therapy is determined by disease stage and the severity of the underlying cirrhosis. The potential positioning of regorafenib by the company in its submission is in those patients who have previously been treated with sorafenib. The company's diagram of current guidelines is reproduced in Figure 2: the ERG has added a red oval showing where

sorafenib is recommended under EASL-EORTC guidelines. In England, sorafenib has recently been reviewed as part of a Cancer Drugs Fund (CDF) reappraisal. The Final Appraisal Determination (FAD) states that: "*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.*"[5] Clinical advice received by the ERG also suggests that sorafenib could also be appropriately used in intermediate stage (B) disease if that disease was not amenable to transarterial chemoembolisation (TACE).

Figure 2: The EASL-EORTC guidelines as represented by the company



CLT=cadaveric liver transplant; DLT=domino liver transplant; HCC=Hepatocellular cancer; mo=months; OS=overall survival; PEI=percutaneous ethanol injection; PST=performance status; RF=radio-frequency ablation; TACE=trans-arterial chemoembolisation

3 CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

3.1 Population

The remit detailed in the final scope issue by the National Institute for Health and Care Excellence (NICE) is to appraise the clinical and cost-effectiveness of regorafenib within its licensed indication for previously treated unresectable HCC. The Summary of Product Characteristics (SmPC) for regorafenib indicates that the therapeutic indication within HCC is for patients "*who have been previously treated with sorafenib.*" A more detailed discussion of the patients in the RESORCE study[6] and those included within the anticipated licence is provided in Section 4.2.1. However, potential key differences are highlighted here. There is uncertainty regarding the generalisability of the results presented to the following groups which were excluded from the RESORCE study:

- adult patients with HCC who were sorafenib intolerant (i.e. having been unable to receive sorafenib at ≥400mg/day for ≥20 of the last 28 days of treatment);
- adult patients with HCC who were Child-Pugh class B;
- adult patients with HCC who had an ECOG PS of 2 or more.

The CS[1] states that regorafenib is "*not recommended for use in patients with severe hepatic impairment (Child-Pugh C) as it has not been studied in this population*".

3.2 Intervention

The intervention evaluated by the company is regorafenib (Stivarga®). Regorafenib is an oral bi-aryl urea that inhibits multiple protein kinases. The standard dose of regorafenib is 160mg daily taken in the form of four 40mg tablets. Within the RESORCE study, two levels of dose reduction due to toxicity were allowed, with reduced doses of either 120mg daily or 80mg daily. The list price for regorafenib is £3,744 per treatment cycle, which consists of three weeks of treatment followed by one week off therapy. The company has agreed a patient access scheme (PAS) with the Department of Health that takes the form of a simple discount ([redacted]): this reduces the cost per treatment cycle to [redacted]. The CS states that the average number of packs received in the RESORCE study was [redacted], equating to an average course of treatment of [redacted] at the list price and [redacted] when the PAS is applied. Any treatment costs accruing beyond the study cut-off date are not included in these estimates. Further details on the intervention are provided in Table 2 of the CS.

3.3 Comparators

The final scope indicated that the sole comparator is best supportive care (BSC). The ERG believes that the RESORCE study, which compared regorafenib in addition to BSC versus placebo in addition to BSC, is an appropriate study to address the decision problem.

3.4 Outcomes

All outcomes listed in the final scope were addressed in the clinical section of the CS. The company's model includes outcomes relating to PFS, OS time to treatment discontinuation and HRQoL (including the impact of AEs).

3.5 Other relevant factors

The company comment that "*the prevalence of liver cancer deaths is higher in socially deprived areas.*" Beyond this statement, this potential equality issue is not considered further within the CS.

4 CLINICAL EFFECTIVENESS

This chapter presents a summary and critique of the company's review of the efficacy and safety of regorafenib (Stivarga®, BAY73-4506) for the treatment of adult patients with HCC who have been previously treated with sorafenib (Nexavar®). The ERG's critique was performed following the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and checklist.[7]

4.1 Critique of the methods of review(s)

The CS[1] reports the methods and results of a systematic review of the efficacy and safety evidence for regorafenib in adult patients with HCC who have been previously treated with sorafenib (see CS,[1] Sections B2.1-B2.13). The systematic review of efficacy and safety evidence was generally well reported. Following a request for clarification from the ERG regarding certain process elements adopted by the company, the ERG considers the review to be generally sound (see company's clarification response,[8] questions A1). There was a single relevant trial: RESORCE. This was a Phase III trial which compared regorafenib with placebo in adult patients with HCC who had previously progressed on sorafenib.

4.1.1 Searches

The company performed one clinical effectiveness search to identify all RCTs investigating the efficacy and safety of regorafenib in previously treated unresectable HCC. For the original searches, several electronic bibliographic databases were searched including MEDLINE [via ProQuest], EMBASE [via ProQuest], the Cochrane Database of Systematic Reviews [via Wiley], the Database of Abstracts of Reviews of Effects [via Wiley], and the Cochrane Central Register of Controlled Trials [via Wiley]. The company searched one clinical trials register (Clinicaltrials.gov via NLM). Conference proceedings websites were searched covering the period from 2014 to January 2017 (American Association for Cancer Research [AACR], American Society of Clinical Oncology [ASCO], Gastrointestinal Cancers Symposium, European Society for Medical Oncology [ESMO], International Liver Cancer Association [ILCA], European Society of Digestive Oncology [ESDO], European Association for the Study of the Liver [EASL], ESMO World Congress on Gastrointestinal Cancer, Japanese Society of Medical Oncology [JSMO], and Chinese Society of Clinical Oncology, and American Association for the Study of Liver Diseases [AASLD].

The company's search strategies were fully reported in CS Appendix D.[9] Since the company searches were completed up until January 2017, the ERG conducted an update search in MEDLINE and EMBASE [via Ovid] on 25th July 2017. A total of 69 records were retrieved from the search. The ERG

found no new studies relevant for the review (see Section 4.2.1) and considers that the company's search strategies were sufficiently comprehensive to retrieve important citations relating to all eligible studies.

4.1.2 Inclusion and exclusion criteria

The inclusion and exclusion criteria for the review of the efficacy and safety of regorafenib are described in CS Appendix D (Table 63, page 187) and are reproduced in Table 3. These criteria describe RCTs measuring the efficacy and safety of regorafenib compared with any intervention, including placebo, in adult patients with HCC who have been previously treated with sorafenib. One RCT satisfied these criteria: RESORCE. This Phase III trial compared regorafenib plus BSC, at a maximum dose of 160mg per day for 3 weeks, followed by a week without treatment, with placebo plus BSC in adult patients with HCC who have previously progressed on sorafenib.

Table 3: Inclusion and exclusion criteria for regorafenib RCTs (reproduced in part from CS, Appendix D, Table 63)

	Inclusion criteria	Exclusion criteria
Population	Adults (aged 18 or older) with advanced HCC	Other (oncology) indications not listed in the inclusion criteria
Intervention	Regorafenib (Stivarga®) (plus BSC)	All interventions not listed in the inclusion criteria
Comparators	Any comparator, including: • BSC* (placebo)	Not applicable
Outcomes	• Overall survival (OS) • Time to progression (TTP) • Progression-free survival (PFS) • Objective tumour Response Rate (ORR) • Disease control • Adverse events (AEs) • Overall AEs • Severe AEs • Quality of Life (QoL) • FACT-Hep • EuroQol – 5 Dimensions (EQ-5D) • Other QoL measurements • All other patient-relevant endpoints	Not applicable
Study design	• Phase II or III randomised controlled trials (RCTs) • Studies published as abstracts, conference presentations or press releases were eligible if adequate data were provided • Systematic reviews or meta-analyses of RCTs**	All other study designs not listed in the inclusion criteria
Language	No language limits	No language limits
*BSC is defined as included any concomitant medications or treatments: antibiotics, analgesics, radiation therapy for pain control (limited to bone metastases), corticosteroids, transfusions, psychotherapy, growth factors, palliative surgery, or any other symptomatic therapy necessary to provide BSC, except other investigational anti-tumour agents or antineoplastic chemo / hormonal / immunotherapy (CS,[1] p.33). **Systematic reviews were eligible for inclusion as a source of references to primary studies		

FACT-Hep - Functional Assessment of Cancer Therapy – Hepatobiliary

4.1.3 Critique of data extraction

The ERG was satisfied that standard systematic review good practice was followed in study selection: relevant papers were independently selected for inclusion at title, abstract and full text stage by two reviewers, with any discrepancies between reviewers resolved through discussion or the intervention of a third reviewer (see CS, Appendix D1.1).

No information was given regarding the data extraction process (for example, the number of reviewers involved, or the nature and extent of any actions taken to minimise error). This was addressed in

response to clarification requests from the ERG, in which the company detailed standard processes for data extraction in systematic reviews (see company's clarification response,[8] question A1). Data extraction was performed by one reviewer and independently checked for errors against the original trial report by a second reviewer. Any discrepancies were resolved through discussion or through the intervention of a third reviewer.

4.1.4 Quality assessment

Critical appraisal of included evidence, using relevant criteria, was performed and reported, although the critical appraisal tool was not identified in the CS (CS,[1] Section B.2.5, Table 18). As with data extraction, no details were provided regarding the critical appraisal process (e.g. number of reviewers undertaking the critical appraisal, processes followed in the event of discrepancies etc.). This was not addressed in response to clarification requests from the ERG: the response focused on data extraction only (see company's clarification response,[8] question A1), so the robustness of the process undertaken is uncertain. However, the identity of the critical appraisal tool used was clarified by the company: this was an adaptation of the Centre for Reviews and Dissemination (CRD) criteria specified in the NICE User Guide (see company's clarification response,[8] question A3).

The CS[1] (page 53) concludes that the RESORCE trial was at 'low risk of bias' across the domains assessed. The ERG also performed a critical appraisal of the relevant RCT to verify the findings reported in the CS. This was conducted by one reviewer (CC) using the Cochrane Risk of Bias tool.[10] The ERG accepts the company's assessments of bias for the domains of selection bias (randomisation, allocation concealment); performance and detection bias (blinding); attrition bias (drop-out, intention-to-treat [ITT] analysis and management of missing data) and reporting bias (this assessment was only confirmed when the company made available the original unpublished protocol: see company's clarification response,[8] question A8). The ERG disagrees with the assessment regarding other types of bias: for example, the extensive role of the funder was acknowledged in the publications, but industry influence is a known potential moderator of outcomes.[11, 12] Overall, however, the ERG assessed the potential risk of bias affecting outcomes in the RESORCE trial to be low. The details of the ERG and CS assessments are provided in Table 4.

Table 4: Risk of bias assessment for the RESORCE trial

Risk of bias	ERG		CS (Appendix D1.3, Table 67)		
Selection bias: Randomisation	"Patients were randomly assigned (2:1) to regorafenib or placebo using a computer-generated randomisation list prepared by the funder. Randomisation was stratified by geographical region (Asia *vs* rest of world), macrovascular invasion (yes *vs* no), extrahepatic disease (yes *vs* no), αfetoprotein concentration (<400 ng/mL *vs* ≥400 ng/mL), and ECOG PS (0 *vs* 1). The randomisation number for each patient was assigned based on information obtained from the interactive voice-response system."[6]	Low	Was randomisation carried out appropriately?	Randomisation was performed via an interactive voice response system (IVRS) using a computer-generated randomisation list. Randomisation was stratified by geographical region (Asia vs. rest of the world), ECOG performance status (0 vs. 1), AFP levels (<400ng/mL vs. ≥400ng/mL), extrahepatic disease (presence vs. absence), and macrovascular invasion (presence vs. absence).	Yes
Selection bias: Allocation concealment		Low	Was the concealment of treatment allocation adequate?	Patients, investigators, and the study sponsor were masked to treatment assignment using the unique randomisation code, assigned via IVRS, which linked them to a treatment arm and specified the treatment assigned. Placebo & active treatments were identical in appearance and given under identical conditions.	Yes
			Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Demographics and baseline disease characteristics were comparable between the regorafenib and the placebo groups.	Yes
Performance bias	"Investigators, patients, and the funder were masked to treatment assignment… Tablets with identical appearance were used for regorafenib and placebo."[6]	Low	Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Investigators received a unique randomisation number for each participant through the IVRS and study drug supply was also managed via IVRS. All patients, investigators, and the study sponsor were masked to treatment assignment through this number. Also, regorafenib and placebo were identical in appearance to preserve blinding.	Yes
Detection bias	Masking of patients and investigators, as outlined above, minimises risk of detection bias for progression and quality of life outcomes. Overall survival (OS) is at very low risk of detection bias. "Investigators were blinded to study treatment for assessment of whether a death was considered related to study drug"[6]	Low			
Attrition bias	All drop-outs and withdrawals were fully reported. Imabalance in withdrawals was due principally to disease progression.	Low	Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	A higher number of patients withdrew from double-blind treatment in the placebo arm of the study (94.3%) than in patients receiving regorafenib (81.5%). The main reason for dropout in both treatment groups was radiological progression.	No

Risk of bias	ERG		CS (Appendix D1.3, Table 67)		
	The primary endpoint (OS) and the secondary endpoints (PFS and TTP) were analysed by ITT. There was no imputatuon of missing data. All patients were analysed in the groups to which they had been randomised.		Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The primary analysis was performed (appropriately) in the FAS (ITT) population. Missing or unevaluable tumour assessments were not used in the calculation of derived efficacy variables unless a new lesion occurred, or the lesions that were evaluated already showed progressive disease (PD). No imputation was performed for missing lesion assessment and tumour response. For example, if a patient missed a scan visit and PD was documented at the next available scan visit, the actual visit date of the first documented PD was used to calculate PFS and TTP. If a date was incomplete, such as only the year and month were available, day 15 of the month was used for the calculation.	Yes / Yes / Yes
Reporting bias	The unpublished trial protocol, provided by the company (see company's clarification response,[8] question A8) permitted a confirmation that all pre-specified outcomes were reported.	Low	Is there any evidence to suggest that the authors measured more outcomes than they reported?	Results of all pre-specified outcomes are reported in full.	No
Other bias	"The funder (Bayer) provided the study drug and worked with the principal investigator (JB) and the study steering committee to design the study. Data collection and interpretation, and preparation of this report, were done by the investigators and the funder. Statistical analyses were performed by the funder... The funder funded writing assistance".[6] "This study was funded by Bayer. Editorial assistance in the preparation of this manuscript was provided by Ann Contijoch (Bayer) and Jennifer Tobin (Choice Healthcare Solutions, with financial support from Bayer)." [6] Authors declare many conflicting interests.[6]	Moderate			

4.1.5 Evidence synthesis

The synthesis for the review of clinical efficacy was a basic descriptive summary of the evidence from the RESORCE trial for the following outcomes: overall survival (OS); progression-free survival (PFS); time to progression (TTP); health-related quality of life (HRQoL) and adverse events (AEs). The CS explains that a meta-analysis was not performed because there was only a single relevant trial (CS,[1] page 72, Section B.2.8) and that an indirect comparison was not performed because the included trial compared the intervention with the most relevant comparator, i.e. BSC/placebo (CS,[1] page 72, Section B.2.9). The ERG accepts these justifications.

4.2 Critique of trials of the technology of interest, their analysis and interpretation

4.2.1 Review of clinical efficacy

The CS provides a detailed description of the RESORCE trial identified by the company satisfying the inclusion criteria, i.e. regorafenib compared with BSC (placebo) in adult patients with HCC who had previously progressed on sorafenib. The CS (Appendix D, page 189, Table 64) identified the following four papers for inclusion: the full trial publication,[6] the protocol (NCT01774344), and two abstracts:LBA-03[13] and LBA28.[13] CS Appendix D (Table 65, pages 190-91) also provided a second list of included studies, which, included the full trial publication, the protocol, the LBA-03 abstract and an earlier Phase II single-arm trial of regorafenib in the relevant population.[14] The CS, Appendix D, Table 65 also erroneously excluded the LBA28 abstract,[13] but reinstated it following a question from the ERG (see company's clarification response,[8] question A2). One full paper[15] and two additional relevant abstracts[6, 16] were identified by the ERG from its own searches. This represented all of the evidence for regorafenib in this population. The Phase II trial, which is included in the CS, was excluded from this report because it does not satisfy the inclusion criteria, which require studies to be comparative.

The trial and its population were slightly different from the NICE scope, which required assessment of regorafenib in all adult patients with HCC who have been previously treated with sorafenib. The RESORCE trial was largely consistent with this population, but did exclude those patients who had discontinued sorafenib treatment on account of toxicity. The inclusion and exclusion criteria for this trial are extensive and are presented in Table 5.

.

Table 5: RESORCE trial inclusion and exclusion criteria (reproduced from CS, Table 6)

Inclusion criteria	Exclusion criteria
• Age ≥ 18 years old • Histologically or cytologically confirmed HCC or non-invasive diagnosis of HCC as per American Association for the Study of Liver Diseases (AASLD) criteria in patients with a confirmed diagnosis of cirrhosis. • Barcelona Clinic Liver Cancer (BCLC) stage Category B or C that could not benefit from treatments of established efficacy with higher priority such as resection, local ablation, chemoembolisation, or systemic sorafenib. • Failure to prior treatment with sorafenib (defined as documented radiological progression per the radiology charter). Randomisation had to be performed within 10 weeks after the last treatment with sorafenib. • Tolerability of prior treatment with sorafenib defined as not less than 20 days at a minimum daily dose of 400 mg QD (every day) within the last 28 days prior to withdrawal. • ECOG PS of 0 or 1 • Child-Pugh status A • Local or loco-regional therapy of intrahepatic tumour lesions (e.g. surgery, radiation therapy, hepatic arterial embolisation, chemoembolisation, radiofrequency ablation, percutaneous ethanol injection, or cryoablation) must have been completed ≥4 weeks before first dose of study medication. Note: patients who received sole intrahepatic intraarterial chemotherapy, without lipiodol or embolising agents were not eligible. • Life expectancy ≥ 3 months • Written consent • At least one uni-dimensional measurable lesion by computed tomography (CT) scan or magnetic resonance imaging (MRI) per Response Evaluation Criteria in Solid Tumors (RECIST) 1.1, and mRECIST for HCC. Tumor lesions situated in a previously irradiated area, or in an area subjected to other loco-regional therapy, may have been considered measurable if there had been demonstrated progression in the lesion.	• Prior liver transplantation or candidates for liver transplantation. • Prior treatment with regorafenib. • Prior and/or concomitant treatment within a clinical study other than with sorafenib during or within 4 weeks of randomisation. • Sorafenib treatment within 2 weeks of randomisation. • Patients with large oesophageal varices at risk of bleeding that were not being treated with conventional medical intervention: beta blockers or endoscopic treatment. • Prior systemic treatment for HCC, except sorafenib. • Permanent discontinuation of prior sorafenib therapy due to sorafenib-related toxicity. • Permanent discontinuation of prior sorafenib therapy due to any cause more than 10 weeks prior to randomisation. • Previous or concurrent cancer distinct from HCC *except* cervical carcinoma in situ, uteri, and/or non-melanoma skin cancer and treated basal cell carcinoma, superficial bladder tumours (Ta, Tis & T1) or any cancer curatively treated > 3 years prior to entry into the study. • Known history or symptomatic metastatic brain or meningeal tumours. • Major surgical procedure or significant traumatic injury within 28 days before randomisation. • Cardiac disease (congestive heart failure > New York Heart Association (NYHA) class 2, cardiac arrhythmias requiring anti-arrhythmic therapy other than beta blockers or digoxin). • Unstable angina (angina symptoms at rest, new-onset angina) or myocardial infarction (MI) within the past 6 months prior to randomisation. • Uncontrolled hypertension (systolic blood pressure [BP] >150 mmHg or diastolic pressure >90 mmHg despite optimal medical management). • Phaeochromocytoma. • Uncontrolled ascites (defined as not easily controlled with diuretic or paracentesis treatment). • Pleural effusion or ascites that caused respiratory compromise (National Cancer Institute [NCI]-common terminology criteria

• Adequate bone marrow, liver and renal function as defined by: haemoglobin >8.5 g/dL; Absolute neutrophil count (ANC) ≥ 1500/mm^3; platelet count ≥ 60,000/mm^3; total bilirubin ≤ 2 mg/dL. Mildly elevated total bilirubin (<6 mg/dL) was allowed if Gilbert's syndrome was documented; alanine aminotransferase (ALT) and aspartate aminotransferase (AST) ≤ 5 X upper limit of normal (ULN); prothrombin time-international normalised ratio (PT-INR) < 2.3 X ULN and partial prothrombin time (PTT) <1.5 X ULN; serum creatinine ≤ 1.5 X ULN; lipase ≤ 2 X ULN; glomerular filtration rate (GFR) ≥30 mL/min/1.73 m^2 per the Modified diet in renal disease (MDRD) study equation. • Women of childbearing potential and men must have agreed to use adequate contraception until at least 2 months for men and for women after the last study drug administration.	for adverse events [CTCAE] Grade ≥2 dyspnoea). • Persistent proteinuria of NCI-CTCAE Grade 3 or higher. Urine dipstick result of 3+ was allowed if protein excretion was < 3.5 g/24 hours. • Ongoing infection > Grade 2 per NCI-CTCAE grading. Hepatitis B was allowed if no active replication was present. Hepatitis C was allowed if no antiviral treatment was required; known history of human immunodeficiency virus (HIV) infection; • Clinically significant bleeding NCI-CTCAE Grade 3 or higher within 30 days before randomisation. • Arterial or venous thrombotic or embolic events such as cerebrovascular accident (including transient ischaemic attacks), deep vein thrombosis or pulmonary embolism within 6 months before the start of study medication. • Unresolved toxicity higher than NCI-CTCAE Grade 1 (excluding alopecia or anaemia) attributed to any prior therapy/procedure. • Any illness or medical condition that was unstable or could have jeopardised the safety of the patient and his/her compliance in the study. • Seizure disorder requiring medication • History of organ allograft; substance abuse, medical, psychological or social conditions that may have interfered with the patient's participation or evaluation of study results; • Inability to swallow oral medications; • Pregnancy or breast-feeding • Non-healing wound, ulcer, or bone fracture. • Renal failure requiring haemo- or peritoneal dialysis. • Known hypersensitivity to any of the study drugs, study drug classes, or excipients in the formulation. • Interstitial lung disease with ongoing signs and symptoms at the time of screening. • Any malabsorption condition. • Close affiliation with the investigational site; e.g. a close relative of the investigator, dependent person of the investigational site that would have had access to study records and electronic case report form [eCRF] data).

HCC - heptaocellular carcinoma; ECOG - Eastern Cooperative Oncology Group; PS - p*erformance status; PT - prothrombin time; INR - International Normalized Ratio; PTT - partial thromboplastin time.*

It is important to note the following groups were excluded from the RESORCE trial:

- adult patients with HCC who were sorafenib intolerant (i.e. having been unable to receive sorafenib at ≥400mg/day for ≥20 of the last 28 days of treatment);
- adult patients with HCC who were Child-Pugh class B;
- adult patients with HCC who had an ECOG PS of 2 or more.

Each of these excluded groups is covered by the BCLC categories B and C (a diagnostic classification that was also included as eligibility criteria for the RESORCE trial) and by the NICE scope.

The Child-Pugh score is an accepted classification of liver function, with higher numbers indicating more impaired liver function and lower numbers (e.g. class A) indicating better preserved liver function.[17] The classification criteria are reproduced in Table 6.

Table 6: Child–Pugh classification[17]

	Score		
Measure	**1 point**	**2 points**	**3 points**
Ascites	Absent	Slight	Moderate
Serum bilirubin (mg/dL)	<2.0	2.0-3.0	>3.0
Serum albumin (g/dL)	>3.5	2.8–3.5	<2.8
Prothrombin time (seconds prolonged)	<4	4-6	>6
Encephalopathy grade	None	1–2	3–4 *

Child–Pugh A: 5 or 6 points; Child–Pugh B: 7–9 points; Child–Pugh C: >9 points

Methodologically, RESORCE is a high quality, international, multicentre placebo-controlled trial (which included five UK centres, four of which had a total of 20 included patients; CS,[1] page 29). Patients (n=573) were randomised 2:1 to receive either oral regorafenib 160mg (4 x 40mg tablets orally once daily) plus BSC (n=379), or 4 x matching placebo tablets plus BSC daily (n=194), for the first 3 weeks of each 4-week cycle. In the fourth week, no study drug/placebo was given. BSC included: antibiotics; analgesics; radiation therapy for pain control (limited to bone metastases); corticosteroids; transfusions; psychotherapy; growth factors; palliative surgery, or any other symptomatic therapy necessary to provide BSC (CS,[1] pages 39-40). A full list of the relative proportions of concomitant medications taken across arms is provided in the clinical study report[18] (CSR), Section 8.7, pages 97-98. The proportion of patients receiving at least one concomitant medication was similar between the two groups (regorafenib 98.2% versus BSC 96.4%[18]). The data cut-off (DCO) for the final analysis was 29th February 2016; the median follow-up was 7.0 months (interquartile range [IQR] 3.7 to 12.6 months[6]). Patients continued masked study treatment until disease progression, death, unacceptable toxicity, substantial non-compliance with the protocol or withdrawal of patient from the study (by physician or patient).

Dose reduction (to 120mg per day or 80mg per day) or interruption was undertaken as required, in response to toxicity or specific AEs (CS,[1] pages 34-38). Doses could be re-escalated once toxicities resolved. When the primary endpoint of the study was reached (i.e. significant survival benefit compared with placebo[6]), patients who were on placebo at that time were offered the opportunity to receive regorafenib through open-label treatment, as long as the risk/benefit profile of regorafenib was positive. Patients were evaluated every cycle for treatment compliance by counting tablets dispensed and returned.

Outcomes and their definitions are described in Table 7. The primary outcome was OS. All disease progression or response outcomes were evaluated by investigators masked to study treatment and based on the RECIST 1.1 criteria and the modified (mRECIST) criteria for HCC regarding the definition of Progressive Disease (PD) and response. The HCC-specific mRECIST[19] is different from RECIST 1.1[20]: it includes amendments developed for the pivotal, sorafenib SHARP trial,[21] requiring cytopathological confirmation of malignancy to classify pleural effusion or ascites as progression, and applies more stringent criteria to define progression due to lymph node involvement at the hepatic hilum or new intrahepatic sites. It also considers complete tumour necrosis on dynamic imaging studies.[6] HRQoL was assessed using two measures: one disease-specific instrument (Functional Assessment of Cancer Therapy – Hepatobiliary: FACT-Hep) and one generic instrument (EQ-5D).

Table 7: Relevant endpoints and measures in the RESORCE trial (adapted from CS, Table 14)

Endpoint	Definition & timing of assessment / measure
Primary endpoint	
Overall survival (OS)	Measured from the date of randomisation until the date of death due to any cause. After the last dose of study medication and the 'end of treatment' visit, all patients entered a follow-up period during which information on survival status was collected.
Secondary endpoints	
Progression-free survival (PFS)	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. Disease progression was based on RECIST 1.1 criteria and the mRECIST criteria for HCC regarding the definition of PD,15 i.e. greater than 20% increase in target lesions. This was performed at screening, every 6 weeks during treatment for the first 8 cycles, and every 12 weeks thereafter.
Time to progression (TTP)	Defined as the time (days) from randomisation to radiological or clinical disease progression.
Objective tumour Response Rate (ORR)	Defined as the proportion of patients with complete response (CR) or partial response (PR) compared with all randomised patients. CR is defined as the absence of all target lesions; PR is defined as a greater than 30% decrease in target lesions. Patients prematurely discontinuing the study without an assessment were considered to be non-responders for the analysis.
Disease Control Rate (DCR)	The rate of subjects, whose best response was not progressive disease compared with all treated subjects (i.e. complete response, partial response or stable disease). In order to be counted as a responder in DCR stable disease had to be maintained for at least 6 weeks. Stable disease is defined as neither PR nor PD.
Tertiary endpoints	
Duration of response	Measured from the date of first documented response (CR or PR) to date of disease progression or death (if death occurred before disease progression).
Duration of stable disease	The time (days) from randomisation to the date that disease progression or death (if death occurred before progression) was first documented. Only calculated for patients who failed to achieve a best response of CR or PR.
Exploratory endpoint	
Overall survival measured from the start of prior sorafenib therapy	Measured from the beginning of prior sorafenib treatment until the date of death due to any cause.
Health-related quality of life (HRQoL):	The FACT-Hep and EQ-5D were both self-administrated by the patient before seeing the physician at baseline, day 1 of each cycle, and at end-of-treatment visit.

Endpoint	Definition & timing of assessment / measure
Primary endpoint	
FACT-Hep (version 4)	FACT-Hep is a 45-item disease-specific module of the FACT questionnaire, used extensively in oncology clinical trials.[22, 23] FACT-Hep consists of five subscales: (1) physical well-being; (2) social/family well-being; (3) emotional well-being; (4) functional well-being; and (5) the hepatobiliary cancer subscale. (1) - (4) are summed to form the FACT-General (FACT-G) total score. (1) - (5) are summed to form the FACT-Hep total score (range 0 to 180).
EuroQol – 5 Dimension (EQ-5D)	The EQ-5D is a generic preference-based quality of life instrument which has been validated in cancer populations to measure both utility and health status.[24] The EQ-5D also contains a visual analogue scale (EQ-visual analogue scale [VAS]), which records the respondent's self-rated health status on a vertical graduated visual analogue scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state).
Other endpoints	
Safety	AE assessment took place at every visit until 30 days after last study treatment (excluding survival assessment). AEs were classified using NCI-CTCAE version 4.03 guidelines

NCI-CTCAE - National Cancer Institute - Common Terminology Criteria for Adverse Events

4.2.2 Results

Participants' baseline characteristics

The baseline demographic and clinical characteristics of patients in the RESORCE trial were comparable across treatment groups (see Table 8). This was the case across almost all characteristics: age; gender; ethnicity; number of target lesions assessed by mRECIST (similar percentages in both treatment groups had two target lesions at baseline, i.e. 46.2% for regorafenib compared with 45.4% for placebo, which was similar when patients were assessed by RECIST 1.1 criteria: [redacted]% for regorafenib compared with [redacted]% for placebo); aetiology (except for alcohol use); ECOG PS; BCLC stage; macroscopic vascular invasion and/or extrahepatic disease; Child-Pugh class; previous local anti-cancer therapy; median time since initial diagnosis of HCC to start of regorafenib treatment; median times since discontinuation of or progression on sorafenib until start of regorafenib treatment. The treatment groups were therefore well-balanced with respect to disease characteristics, prognostic factors[25] and progression on sorafenib. Alcohol use was reported as aetiology for 23.8% in the regorafenib group and 28.4% in the placebo group. There was a difference in median alpha-fetoprotein (AFP) between arms: 183.2ng/ml (range 1.0-477591.0ng/ml) in the regorafenib arm compared with 234ng/ml (range 1.0-310229.1ng/ml) in the placebo arm. However, this was less noticeable when categorised as < or ≥400ng/ml; clinical advice received by the ERG confirmed that this categorisation was appropriate.

It is important to note that patients in the RESORCE trial were exclusively diagnosed as ECOG PS 0 and 1 (100% in each arm) and were almost exclusively diagnosed as Child-Pugh class A (98% in the regorafenib arm compared with 97% in the placebo arm). Patients who were Child-Pugh class B were excluded from the trial, but the status of a very small number of patients changed between recruitment and first administration of the study drug. The RESORCE trial therefore does not provide any meaningful evidence on patients who are BCLC stage B or C who are also ECOG PS 2 or Child-Pugh class B (two populations which are covered by the marketing authorisation of regorafenib for this indication). This limitation is acknowledged by the company (CS[1] page 85 and clarification response,[8] question A6).

The CS[1] (Section B.2.13, pages 86-89) also presents the findings of a retrospective audit of medical records by King *et al* (2017) reporting details of 484 sorafenib-treated patients in 15 hospitals in the UK between 2007 and 2013.[26] Where data are available and a comparison is possible, the baseline demographics and clinical characteristics of these patients were generally similar to the RESORCE trial. However, the patients in the sorafenib audit are older (mean age 68 years compared to 61 or 62 years in the RESORCE trial); they are much less likely to have HCC caused by hepatitis B (12.3% in the sorafenib audit compared with 38% in the RESORCE trial) and are much less likely to have extrahepatic disease (40% in the sorafenib audit compared with between 70% and 76% in the RESORCE trial). The other principal differences are the considerably higher proportions of patients with Child-Pugh class B (21% in the sorafenib audit compared with 0% in the RESORCE trial) and ECOG PS 2 (16% in the sorafenib audit compared with between 1% and 3% in the RESORCE trial). These patients are currently covered by the license for regorafenib.

Table 8: Patient baseline characteristics in the RESORCE trial

	Regorafenib N=379 (%)	**Placebo N=194 (%)**	**Sorafenib audit N=484 (%)**
Age (yr) (mean ± SD)	61.8 ±12.4	61.1±11.6	68
Median age (range)	64 (54-71)	62 (55-68)	
< 65 years	199 (52.5)	116 (59.8)	
≥ 65 years	180 (47.5)	78 (40.2)	
Sex – no. (%)			
Male	333 (87.9)	171 (88.1)	325 (72.5)
Female	46 (12.1)	23 (11.9)	66 (14.7)
Not reported			57 (12.7)
Race			
White	138 (36.4)	68 (35.1)	
Black	6 (1.6)	2 (1.0)	
Asian	156 (41.2)	78 (40.2)	
White / Black	2 (0.5)	1 (0.5)	
Not reported	77 (20.3)	45 (23.2)	
Region – no. (%)			
Asia	143 (37.7)	73 (37.6)	
Rest of World	236 (62.3)	121 (62.4)	

	Regorafenib N=379 (%)	**Placebo N=194 (%)**	**Sorafenib audit N=484 (%)**
Number of target lesions (mRECIST) n=372			
1	67 (17.7)	31 (16.0)	
2	175 (46.2)	88 (45.4)	
3	68 (17.9)	37 (19.1)	
4	43 (11.4)	26 (13.4)	
5	19 (5.0)	12 (6.2)	
Cause of disease (Aetiology)* – no. (%)			
Hepatitis C	78 (20.6)	41 (21.1)	70 (15.6)
Alcohol use	90 (23.8)	55 (28.4)	110 (24.6)
Hepatitis B	143 (37.7)	73 (37.6)	55 (12.3)
Genetic / metabolic	16 (4.2)	6 (3.1)	
Non-Alcoholic steatohepatitis	25 (6.6)	13 (6.7)	
Unknown	66 (17.4)	32 (16.5)	
Other	12 (3.2)	4 (2.1)	
ECOG performance status – no. (%)			
0	247 (65)	130 (67)	117 (26.1)
1	132 (35)	64 (33)	218 (48.7)
2			94 (21.0)
3			6 (1.3)
No data			13 (2.9)
BCLC stage - no. (%)			
A (early)	1 (0.3)	0	3 (0.7)
B (intermediate)	53 (14.0)	22 (11.3)	104 (23.2)
C (advanced)	325 (85.8)	172 (88.7)	322 (71.9)
No data			19 (4.2)
Macroscopic vascular invasion – no. (%)			
Yes	110 (29.0)	54 (27.8)	91 (20.3)**
No	269 (71.0)	140 (72.2)	161 (35.9)**
No data			196 (43.8)
Extrahepatic disease – no. (%)			
Yes	265 (69.9)	147 (75.8)	172 (38.4)
No	114 (30.1)	47 (24.2)	269 (60.0)
No data			7 (1.6)
Macroscopic vascular invasion and/or extrahepatic disease – no. (%)	304 (80)	162 (84)	
Child-Pugh class – no (%)			
A	373 (98.4)	188 (96.9)	343 (76.6)
B†	5 (1.3)	6 (3.1)	72 (16.1)
C‡			2 (0.4)
No data			31 (6.9)
Child-Pugh score – no (%)			
5	244 (64.4)	118 (60.8)	
6	129 (34.0)	70 (36.1)	
7†	5 (1.3)	5(2.6)	
8	0	1 (0.5)	
AFP (ng/ml)			

	Regorafenib N=379 (%)	Placebo N=194 (%)	Sorafenib audit N=484 (%)
Mean (± S.D.)	13507.9 (±49056.8)	12621.7 (±38472.3)	
median (range)	183.2 (1.0-477591.0)	234 (1.0-310229.1)	
<400 ng/mL	217 (57.3)	107 (55.2)	227 (50.7)
≥400 ng/mL	162 (42.7)	87 (44.9)	141 (31.5)
Previous therapy – no. (%)			
Local anti-cancer therapy	256 (67.9)§	133 (68.6)	
Including use of drug given locally	224 (59.1)	115 (59.3)	
Radiotherapy	48 (12.7)	37 (19.1)	
Systemic anticancer therapy	379 (100)	194 (100)	
Time from initial HCC diagnosis to start of regorafenib treatment – (months)			
Median (IQR)	21 (11-38)	20 (12-32)	
Mean (SD)	29 (28)	27 (22)	
Duration of sorafenib treatment (months)			
Median (IQR)	7.8 (4.2-14.5)	7.8 (4.4-14.7)	
Time from progression on sorafenib to start on regorafenib			
Median (IQR)	1.4 (0.9-2.3)	1.4 (0.9-2.2)	
Time from discontinuation of sorafenib to start on regorafenib			
Median (IQR)	0.9 (0.7–1.3)	0.9 (0.7–1.3)	

Data reproduced from CS, Table 15, pages 49-50, Table 30, pages 87-88; CSR Table8-5, and Table 1[6]; and the sorafenib audit study[26].
SD - standard deviation; HCC - hepatocellular carcinoma; IQR - interquartile range:
** Patients may have had more than one aetiology of HCC ** reports vascular invasion*
‡Regorafenib is not licenced for Child-Pugh class C [redacted]
† The information in this table is based on the last observations on or before the first study drug intake. Changes may have occurred between the screening of patients and their first day of study drug intake. During the study, it was found that 3 patients were on anticoagulant medication which, per the study protocol, led to Child-Pugh classification of B.

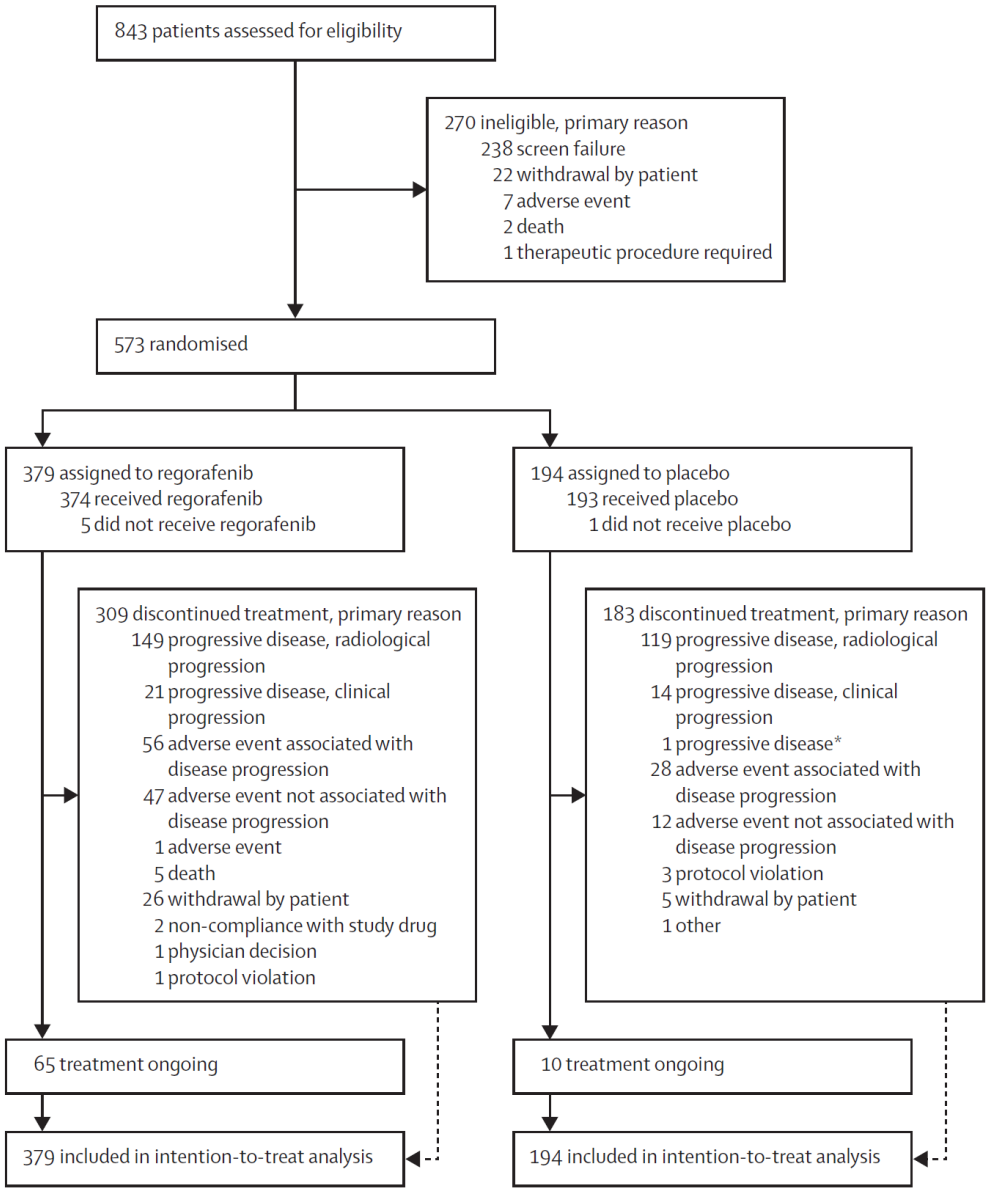
Participant flow and numbers

A total of 573 eligible patients were randomised 2:1 to regorafenib (n=379) and placebo (n=194), but 567 started treatment (five patients in the regorafenib group and one in the placebo group withdrew before first administration of the study drug). The reasons for these withdrawals were not provided in the CSR[18] (Section 8.2, page 82) but were provided by the company in response to a request from the ERG; they were principally due to the erroneous inclusion of patients who did not satisfy the eligibility criteria (see company's clarification response,[8] question A5). The patient in the placebo arm was excluded due to becoming Child-Pugh class B between randomisation and first study treatment (and therefore no longer satisfied the Child-Pugh class A inclusion criterion), but at least some of the 11 other patients who experienced a similar change in status before first study treatment were still included (see Table 8†). This anomaly is not explained. The ITT efficacy analysis included all randomised

patients (n=573), whilst the safety analysis only included patients who had started treatment (n=567). Details of the participant flow through the trial and reasons for discontinuation are provided in Figure 3. Three hundred and nine (83%) of the regorafenib patients who started treatment on regorafenib discontinued treatment, compared with 183 (95%) in the placebo arm of the trial. The numbers discontinuing due to disease progression were 226 (60%) in the regorafenib group and 162 (84%) in the placebo group. Discontinuations due to AEs not associated with disease progression were 15% (47/309) in the regorafenib arm, compared with 7% (12/183) in the placebo arm.

During the double-blind period (before reaching the primary endpoint), the median treatment duration for patients assigned to receive regorafenib was 3.6 months (IQR 1.6-7.6 months) compared with 1.9 months (IQR 1.4-3.9 months) for patients assigned to placebo (CS,[1] page 39). The median daily dose during the double-blind treatment period was reported to be 159.3mg for regorafenib-treated patients (CS,[1] page 39). The mean daily dose of regorafenib was 144.1mg (standard deviation [SD] 21.3mg) and 157.4mg of placebo (SD 10.3mg). Excluding treatment delays or interruptions, almost half of the regorafenib group (184 of 374 [49%]) received the full protocol dose (160mg/day) with no reductions (CS,[1] page 39). Full details of concomitant and disallowed concomitant medications, and required therapeutic and diagnostic and therapeutic procedures, were also provided in the CS (see CS,[1] pages 39-42).

Figure 3: Participant flow in the RESORCE trial[6]



Primary outcome

4.2.2.1 Overall survival

In the RESORCE trial, median OS was reported to be 10.6 months (95% confidence interval (CI) 9.1-12.1 months) in patients randomised to regorafenib compared with 7.8 months (95% CI 6.3-8.8 months) in patients randomised to placebo. The estimated hazard ratio (HR) for OS for regorafenib compared with placebo was 0.63, 95% CI 0.50-0.79, one-sided p=0.000020 (previously published as 0.62, 95% CI 0.50-0.78, p<0.001[13]). This represents a statistically significant reduced risk of death of 37% in the regorafenib group compared with the placebo group. This satisfies the primary objective of the trial in terms of an HR of 0.7 or better, but not the targeted improvement of 43% increase in median OS

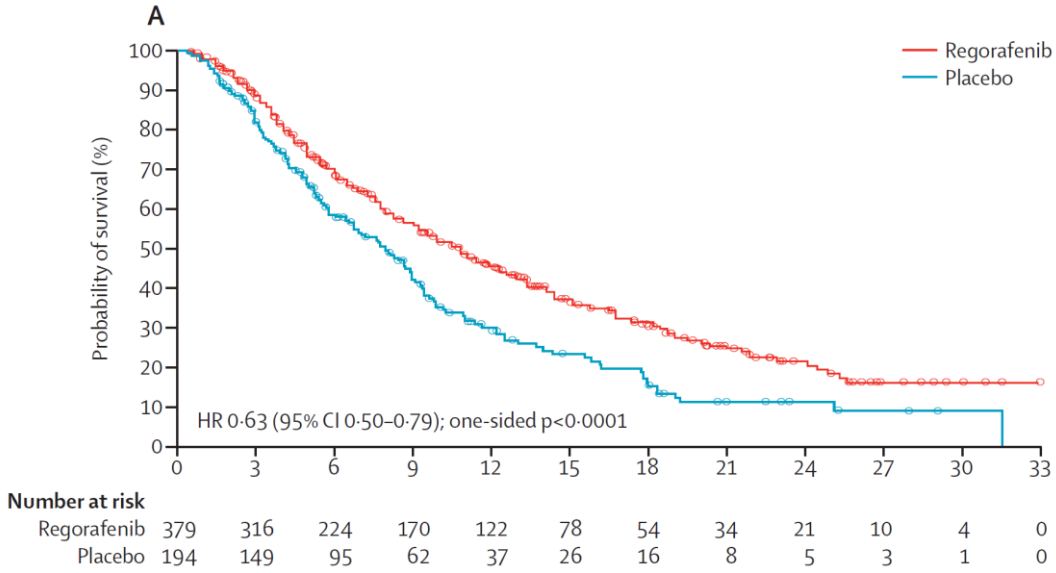
compared to placebo ([redacted]) (see CS,[1] Table 17, page 51). Details are presented in Table 9 and the Kaplan-Meier curve is reproduced in Figure 4**.**

Table 9: Analyses of overall survival in the RESORCE study (FAS; mRECIST) (reproduced from CS, Table 19)

	Regorafenib (N=379)	**Placebo (N=194)**
Number of patients (%) with event	[redacted]	[redacted]
Number of patients (%) censored	[redacted]	[redacted]
Median overall survival, days (95% CI), Range (without censored values)	[redacted]	[redacted]
Median overall survival, months (95% CI), Range (without censored values)	[redacted]	[redacted]
Primary analysis		
Hazard ratio[a]: Stratified IVRS	[redacted]	
95% CI for hazard ratio:	[redacted]	
p-value (one-sided) from log-rank test)	0.000020	

CI - confidence interval; FAS - full analysis set; IVRS - interactive voice response system
[a] An HR <1 indicates superiority of regorafenib 160mg over placebo.
Median, percentile and other 95% CIs computed using Kaplan-Meier estimates. HR and its 95% CI based on a stratified (IVRS) Cox regression model.
Durations were manually converted from days to months (1 month=30.44 days)

Figure 4: Kaplan-Meier Curve for OS (FAS; mRECIST) (reproduced from Bruix *et al*, 2017, Figure 2A[6])



Secondary outcomes

4.2.2.2 Progression-free survival (PFS)

In the RESORCE trial, median PFS as measured by mRECIST was statistically significantly better for regorafenib (3.1 months, 95% CI 2.8–4.2 months) than for placebo (1.5 months, 95% CI 1.4–1.6 months): HR, 0.46, 95% CI 0.37-0.56; p<0.0001. This represents a 54% reduced risk of progression for regorafenib group compared with placebo. Details are presented in

Table **10** and Figure 5.

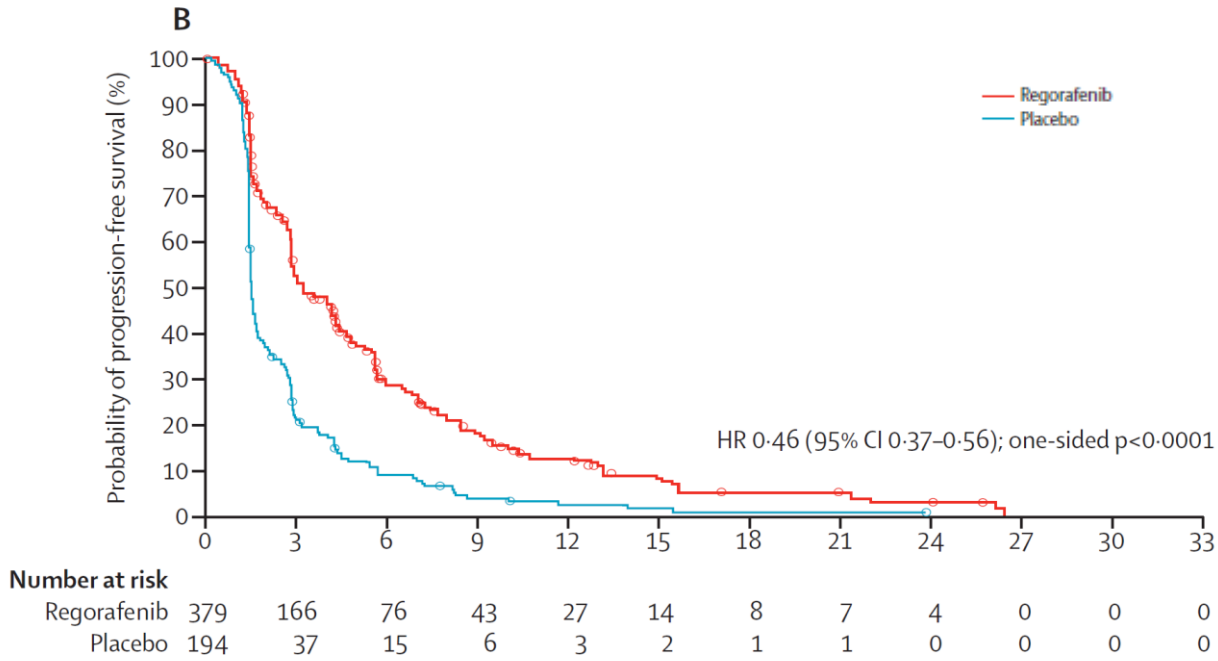
Table 10: Analyses of PFS in the RESORCE study (FAS; mRECIST) (reproduced from CS, Table 20)

	Regorafenib (N=379)	**Placebo (N=194)**
Number of patients (%) with event	[redacted]	[redacted]
Number of patients (%) censored	[redacted]	[redacted]
Median PFS, days (95% CI), Range (without censored values)	[redacted]	[redacted]
Median PFS, months (95% CI), Range (without censored values)	[redacted]	[redacted]
Primary analysis		
Hazard ratio[a]: Stratified IVRS	[redacted]	

95% CI for hazard ratio	[redacted]
p-value (one-sided) from log-rank test) [b]	<0.0001

CI - confidence interval; FAS - full analysis set; IVRS - interactive voice response system
[a] An HR <1 indicates superiority of regorafenib 160mg over placebo.
Median, percentile and other 95% CIs computed using Kaplan-Meier estimates. HR and its 95% CI based on a stratified (IVRS) Cox regression model. Durations were manually converted from days to months (1 month=30.44 days)

Figure 5: Kaplan-Meier curve for PFS in the RESORCE study (FAS; mRECIST) (reproduced from Bruix et al, 2017, Figure 2B[6])



As measured by RECIST 1.1, median PFS was [redacted] months (95% CI [redacted] months) for regorafenib compared with [redacted] months (95% CI [redacted] months) for placebo: HR 0.43, 95% CI 0.35–0.52; one-sided $p<0{\cdot}0001$.[6]

4.2.2.3 Time to progression (TTP)

In the RESORCE trial, median TTP as measured by mRECIST was statistically significantly better for regorafenib (3.2 months, 95% CI 2.9–4.2 months) than for placebo (1.5 months, 95% CI 1.4–1.6 months): HR, 0.44, 95% CI 0.36-0.55; $p<0.0001$. This represents a 56% reduced risk in TTP in the regorafenib group compared with the placebo group. Details are presented in Table 11 and Figure 6.

As measured by RECIST 1.1, median TTP (95% CI) was 3.9 months for regorafenib (95% CI 2.9–4.2 months) compared with 1.5 months for placebo (95% CI 1.4–1.6 months): HR, 0.41, 95% CI 0.34-0.51; $p<0.0001$.[6]

Table 11: Analyses of TTP in the RESORCE study (FAS; mRECIST) (reproduced from CS, Table 21)

	Regorafenib	**Placebo**

	(N=379)	(N=194)
Number of patients (%) with event	[redacted]	[redacted]
Number of patients (%) censored	[redacted]	[redacted]
Median TTP, days (95% CI), Range (without censored values)	[redacted]	[redacted]
Median TTP, months (95% CI), Range (without censored values)	[redacted]	[redacted]
Primary analysis		
Hazard ratio [a]: Stratified IVRS	[redacted]	
95% CI for hazard ratio:	[redacted]	
p-value (one-sided) from log-rank test) [b]:	<0.0001	

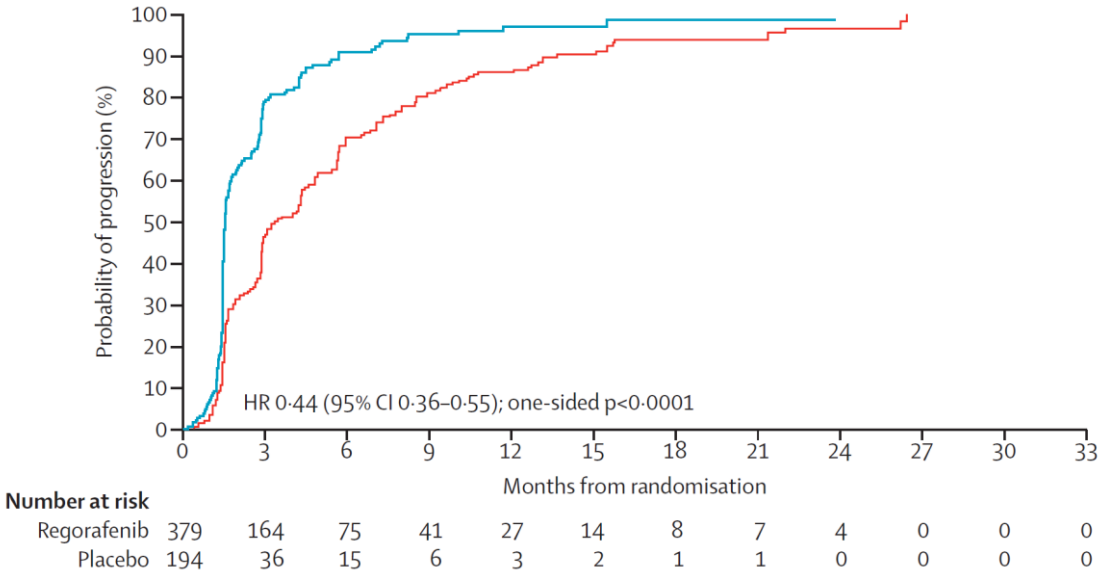
CI - confidence interval; FAS - full analysis set; IVRS - interactive voice response system

[a] *An HR <1 indicates superiority of regorafenib 160mg over placebo.*

Median, percentile and other 95% CIs computed using Kaplan-Meier estimates. HR and its 95% CI based on a stratified (IVRS) Cox regression model.

Durations had been manually converted from days to months (1 month=30.44 days)

Figure 6: KM estimates of the TTP rate during RESORCE (FAS; mRECIST) (reproduced from Bruix *et al* 2017, Figure 2C[6])



4.2.2.4 Response

The objective response rate (ORR), the aggregation of CR and PR, according to mRECIST, was statistically significantly higher in the regorafenib group than in the placebo group (11% compared with 4%; *p*=0.0047, see Table 12**.**

Table 12: Response to therapy in the RESORCE study (FAS; mRECIST) (reproduced from CS, Table 22 and Bruix *et al*)

Best overall response	**Regorafenib N=379 (100%) [95% CI]**	**Placebo N=194 (100%) [95% CI]**
Complete response (CR)	2 (1%) [<1%; 2%]	0
Partial response (PR)	38 (10.0%) [7%; 14%]	8 (4%) [2%; 8%]
Stable disease (SD)	206 (54%) [49%; 59%]	62 (32%) [26%; 39%]
Non-CR / Non-PD	1 (0.3%) [0.0%; 1.5%]	0
Progressive disease (PD)	86 (23%) [19%; 27%]	108 (56%) [48%; 63%]
Not evaluable (NE)	19 (5%) [3%; 8%]	8 (4%) [2%; 8%]
Not assessed (NA)	27 (7%) [5%; 10%]	8 (4%) [2%; 8%]
Clinical progression	86 (23%) [19%; 27%]	40 (21%) [15%; 27%]
Response Rate	**40 (11%)**	**8 (4%)**
Disease Control Rate	**247 (65%)**	**70 (36%)**
Comparison of treatments – Inferential Statistics		

Regorafenib versus placebo	**Difference**	**[95% CI]**	***p*-value**
Response rate	-6.61	[-10.84, -2.39]	0.0047
Disease control rate	-29.31	[-37.52, -21.11]	<0.0001

CI - confidence interval; CR - complete response; FAS - full analysis set; HCC - hepatocellular carcinoma; mRECIST - modified RECIST for HCC; N - number of patients; NA - not assessed; NE - not evaluable; PD - progressive disease; PR - partial response; SD - stable disease

Based on mRECIST, the disease control rate (DCR), a combination of CR + PR +Stable Disease (SD), was also statistically significantly higher in the regorafenib group compared with the placebo group (65% compared with 36%; $p<0.0001$). Stable disease is defined as neither PR nor PD. Using the mRECIST criteria, two patients were reported as having had a complete response (CR) (0.5%) in the regorafenib arm (compared with no patients in the placebo arm, see Table 12**.**

It should be noted that according to RECIST 1.1 criteria, no patients achieved CR and the overall response rate was reduced: [redacted] vs [redacted] (regorafenib vs placebo), [redacted] (Online Table 6[6]), compared with 11% vs 4%, respectively, $p<0.0001$, according to mRECIST (see Table 12).

In terms of the tertiary endpoints, based on mRECIST criteria, the CS reported that the median duration of response and median duration of stable disease were longer in the regorafenib group than in the placebo group, however these differences were not statistically significant (no *p*-values were given, see Table 13).

Table 13: Duration of response and stable disease (FAS; mRECIST) (reproduced from CS, Tables 24 and 25)

Duration of response	**Regorafenib (N=40)**	**Placebo (N=8)**
Number of patients (%) with event	30 (75.0%)	5 (62.5%)
Number of patients (%) censored	10 (25.0%)	3 (37.5%)
Median [95% CI], months	3.5 (1.9-4.5)	2.7 (1.9, NE)
Duarion of stable disease	**Regorafenib (N=206)**	**Placebo (N=62)**
Number of patients (%) with event	151 (73.3%)	56 (90.3%)
Number of patients (%) censored	55 (26.7%)	6 (9.7%)
Median (95% CI) in months	5.5 (4.3 – 5.6)	3.1 (2.8, 4.2)

CI - confidence interval; CR - complete response; FAS - full analysis set

4.2.2.5 Health-related quality of life (HRQoL)

The CS[1] (page 67) reports that 'more than 80% of regorafenib and placebo patients completed questionnaires' and that, 'Of these, approximately 90% in either treatment group were valid for analyses'. The CSR[18] (Sections 9.3.3.3.1 and 9.3.3.3.2) refers to these figures, which are otherwise unpublished. The trial found that quality of life scores were generally similar across arms (see Table 14), but all of the different measures consistently favoured placebo compared with regorafenib (CS,[1] pages 67-70 and Bruix *et al* 2017[6, 13]). The Least-Squares Mean (LSM) time-adjusted Area Under the Curve (AUC) analysis found that only two measures produced a statistically significant difference between arms: the FACT-Hep total and Trial Outcome Index (TOI, a subscale of FACT-Hep) both favoured placebo compared with regorafenib ($p<0.0001$ and $p=0.0006$, respectively). The trial publications[6, 13] and the CS (page 67) stated that even though the differences were statistically

significant, they were not clinically meaningful because they did not exceed minimally important thresholds for the differences, as established in the literature (a change of 8-9 points for FACT-Hep[27] and 7-8 points, for the EQ-5D VAS[24]).

Table 14 : Summary of patient-reported outcomes; LSM time-adjusted AUC (FAS) (reproduced from CS, Table 27, page 68 and Bruix *et al,* 2017[6, 13])

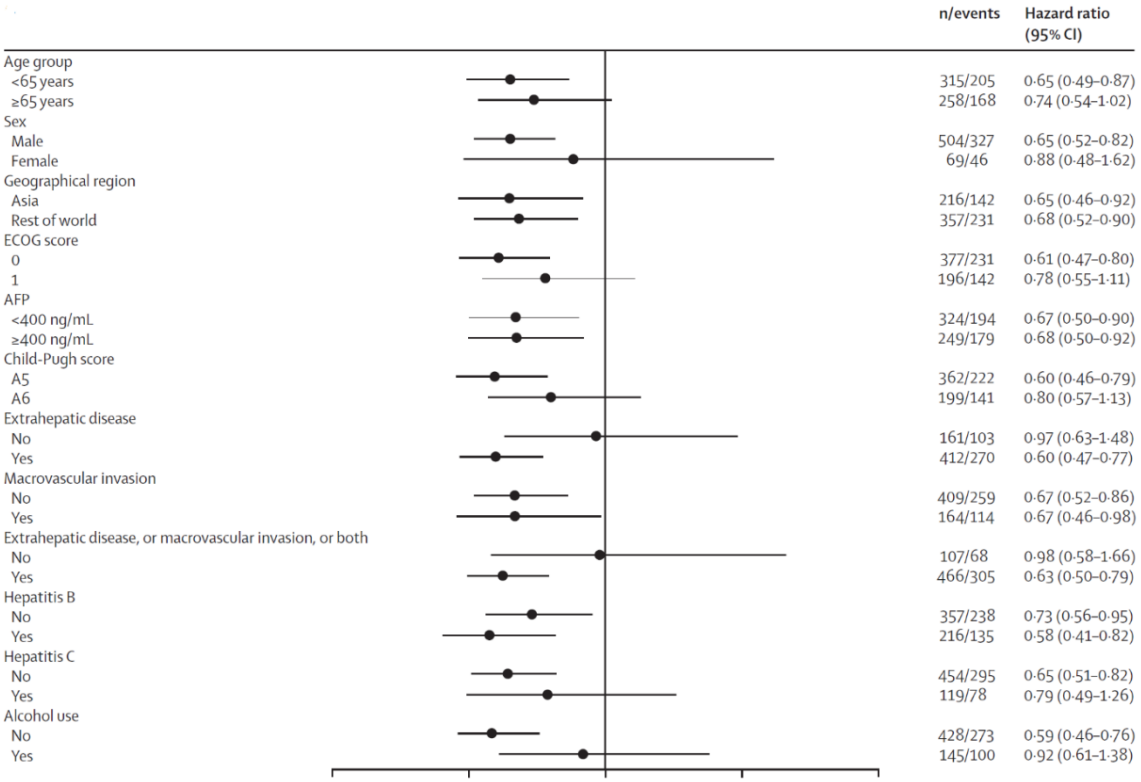
LSM time-adjusted AUC (95% CI)	Regorafenib	Placebo	Difference	*p*-value	MID
EQ-5D index	0.76 (0.75, 0.78)	0.77 (0.75, 0.79)	-0.01 (-0.03, 0.02)	0.4695	0.1
EQ-5D VAS	71.68 (70.46, 72.90)	73.45 (71.84, 75.06)	-1.77 (-3.58, 0.04)	0.0558	10
FACT-G	75.14 (74.12, 76.16)	76.55 (75.20, 77.90)	-1.41 (-2.93, 0.11)	0.0698	6-7
FACT-Hep total	129.31 (127.84, 130.79)	133.17 (131.21, 135.12)	-3.85 (-6.06, -1.65)	0.0006	8-9
Trial outcome index	91.47 (90.30, 92.64)	95.52 (93.98, 97.07)	-4.05 (-5.79, -2.31)	<0.0001	7-8

AUC - area under curve; FACT - Functional Assessment of Cancer Therapy; FACT-G - FACT-General; FACT-Hep - FACT-Hepatobiliary; LSM - Least squares mean; MID - minimally important difference; VAS - visual analogue scale

4.2.2.6 Subgroup and exploratory analyses

Subgroup analyses were conducted for OS, PFS and TTP. Full details of all of the subgroup analyses are provided in the CS, Appendix E, but the forest plot for the OS subgroup analyses is reproduced here. All of these analyses demonstrated consistent benefit for patients treated with regorafenib, regardless of geographical location, age, gender, AFP, aetiology or other covariates (see Figure 7**)**. A published abstract also reported that, while there was a consistent OS benefit regardless of pattern of progression under sorafenib, patients had a substantially worse prognosis if they developed new extrahepatic lesions under previous sorafenib treatment: on regorafenib, 9.7 months with new extrahepatic lesions compared with 14.7 months with no new such lesions; compared with 8.2 months and 10.5 months respectively on placebo.[6] A subgroup analysis of Chinese patients reported results similar to the overall trial, albeit being a younger population with slightly shorter survival times.[16] The CS correctly acknowledged (Appendix E) not only that the RESORCE trial was not powered for subgroup analyses, but also that the number of patients in some subgroups was small, with low event rates. This means that the results of these analyses should be interpreted with caution (see CS Appendix E[9]).

Figure 7: Forest plot of subgroup analyses – overall survival (FAS) (reproduced from Bruix *et al*, 2017, Figure 3A[6])



Note: the left-hand side of the line of no effect favours regorafenib

Sensitivity analyses

Sensitivity analyses were conducted for the outcomes of OS, PFS and TTP to take into account any differences between the primary analysis stratification data collected by the investigator at the time of randomisation (the IVRS analysis), and those collected later on each patient's Case Report Form (CRF) by a validated electronic system for data collection (the RAVE analysis), as well as an analysis that did not use the stratification factors (CS,[1] page 57). The findings across these sensitivity analyses were consistent with the primary analysis using the data according to the IVRS (see CS,[1] pages 56-61).

Exploratory analysis

An exploratory analysis evaluating OS from the beginning of previous sorafenib treatment was also undertaken for the RESORCE trial. This demonstrated that the median OS was statistically significantly improved by the sequence of sorafenib followed by regorafenib from [redacted] months on placebo (95% CI [redacted] months) to [redacted] months on regorafenib (95% CI [redacted] months) (a difference of [redacted] months, see

Table 15).

Table 15: OS from start of prior sorafenib treatment (adapted from CS, Table 26)

		Regorafenib (N=379)	Placebo (N=194)
	n N missing	[redacted]	[redacted]
Time (days) from start of sorafenib to progression while on sorafenib	Median (95% CI) (range)	[redacted]	[redacted]
Time (days) from start of sorafenib to progression on regorafenib	Median (95% CI) (range)	[redacted]	[redacted]
Time (days) from start of sorafenib to death	Median (95% CI) (range)	[redacted]	[redacted]

CI - confidence interval

4.2.3 Safety

AEs were assessed using the MedDRA preferred terms (https://www.meddra.org/) and NCI-CTCAE grading (https://evs.nci.nih.gov/ftp1/CTCAE/About.html) for the safety population (n=567). It should be noted that although many data were similar or the same, the data on AEs presented below (as published[6]) differ from some of the data presented (as academic-in-confidence) in the CS (pages 74-77 and Table 28) and in the CSR. The CS (Section B.2.10, pages 74-80) and the CSR (Section 10.3.2) AE rates from the 'safety analysis set', are almost always lower. However, in such cases, the differences in comparison with placebo are either similar or indicate a comparable difference between arms, e.g. anaemia of any grade [redacted] compared with 16% and 11% (published); for fever, [redacted] compared with 19% vs 11% (published); and for fatigue, [redacted] compared with 40% vs 32% (published, see Table 16). Given these disparities, the ERG has decided to present only the published data. Similar discrepancies exist for the treatment-emergent drug-related AEs.

AEs were frequent (all patients receiving the study drug experienced at least one AE) (see Table 16) and 93% of regorafenib patients experienced treatment-emergent drug-related AEs compared with 52% of placebo patients. All common AEs were much more frequent in the regorafenib group than in the placebo group. The principal AEs were: hand foot skin reaction (53% in the regorafenib arm compared with 8% in the placebo arm); diarrhoea (41% vs 15%); fatigue (40% vs 32%); hypertension (41% vs 6%); and anorexia (31% vs 15%). The frequency of the most common AEs was consistent with those in the Phase II trial[14], with the exception of hypothyroidism, which occurred in 15% of regorafenib patients in the Phase II trial, but only 6.4% of regorafenib patients in the RESORCE trial (CSR, Table 10-3, page 149). The relative frequency of other events was more consistent.

Table 16: **Incidence of any adverse event with a frequency of ≥10% and ≥5% difference between regorafenib and placebo[6]**

	Treatment-emergent (%)		Treatment-emergent drug-related (%)	
Adverse event	**Regorafenib (n=374)**	**Placebo (n=193)**	**Regorafenib (n=374)**	**Placebo (n=193)**
Any	100	93	93	52
Hand foot skin reaction	53	8	52	7
Diarrhoea	41	15	33	9
Fatigue	40	32	29	19
Hypertension	31	6	23	5
Anorexia	31	15	24	6
Increased blood bilirubin	29	18	19	4
Abdominal pain	28	12	9	3
Increased AST	25	20	13	8
Fever	19	7	4	2
Constipation	17	11	6	2
Anaemia	16	11	6	1
Hypoalbuminaemia	15	8	2	0
Weight loss	14	5	7	2
Oral mucositis	13	3	11	3
Vomiting	13	7	7	3
Thrombocytopenia	10	3	5	1
Hypophosphataemia	10	2	6	1
Hoarseness	10	1	9	0

Rates of AEs of Grade 3 or higher were reported as 79.9% in the regorafenib group compared with 58.5% in the placebo group.[13] More regorafenib patients than placebo patients experienced Grade 3 (46% compared with 16%) and Grade 4 (4% compared with 1%) treatment-emergent drug-related AEs. Some Grade 3 and 4 AEs were also much more frequent in the regorafenib group than in the placebo group (see Table 17). The principal Grade 3 AEs were: hand foot skin reaction (13% in the regorafenib arm compared with 1% in the placebo arm); hypertension (15% vs 5%); increased blood bilirubin (10% vs 8%); increased aspartate aminotransferase (AST, 10% vs 10%); fatigue (9% vs 5%); and hypophosphataemia (8% vs 2%). The only Grade 4 AEs affecting more than 1% of patients all occurred in the regorafenib arm: these were increased blood bilirubin (3%), increased alanine aminotransferase (ALT, 3%) and increased AST (2%).

According to the principal trial publication,[6] serious adverse events (SAEs) occurred in 166 (44%) patients in the regorafenib group and 90 (47%) patients in the placebo group. These SAEs were attributed to the study drug in 39 (10%) regorafenib patients and five (3%) placebo patients.[6] According to the CS[1] (page 79), drug-related treatment-emergent severe AEs (TESAEs) were relatively low in both groups, but higher in regorafenib-treated patients compared with those receiving placebo (10% [n=39] vs. 3% [n=5]). The most common TESAEs (>2%) were general physical health deterioration

([redacted]% in regorafenib patients compared with [redacted]% in placebo patients); ascites ([redacted]%; vs [redacted]%) and hepatic failure ([redacted]% vs [redacted]%).

According to the SmPC,[28] regorafenib has been associated with an increased incidence of haemorrhagic events, which were mostly mild to moderate, but some of which were fatal. As a result, close monitoring is recommended for patients who are predisposed to bleeding. In the RESORCE trial, according to the CS[1] (page 78, Table 29) and Online Table 11,[6] the incidence of haemorrhage events of ≥ Grade 3 was higher in the placebo group (15 patients=8%) than the regorafenib group (21 patients=6%), but the incidence of drug-related haemorrhage events of ≥ Grade 3 was higher in the regorafenib group (6 patients=1.6%) than the placebo group (0 patients).

Table 17: **Incidence of Grade 3 and 4 adverse events (≥2%)[6]**

Adverse event	Treatment-emergent (%)				Treatment-emergent drug-related (%)			
	Regorafenib (n=374)		Placebo (n=193)		Regorafenib (n=374)		Placebo (n=193)	
	Grade 3	Grade 4	Grade 3	Grade 4	Grade 3	Grade 4	Grade 3	Grade 4
Any	56	11	32	7	46	4	16	-
Hand foot skin reaction	13	-	-	-	13	-	-	-
Diarrhoea	3	-	-	-	2	-	-	-
Fatigue	9	-	5	-	6	-	2	-
Hypertension	15	-	5	-	13	-	3	-
Anorexia	3	-	2	-	3	-	-	-
Increased blood bilirubin	10	-	8	3	6	-	2	-
Abdominal pain	3	-	4	-	-	-	-	-
Increased AST	10	-	10	2	4	-	5	-
Ascites	4	-	6	-	-	-	-	-
Anaemia	4	-	5	-	-	-	-	-
Increased ALT	3	-	11	3	2	-	4	-
Hypoalbuminaemia	2	-	8	-	-	-	-	-
Weight loss	2	-	5	-	-	-	-	-
Back pain	2	-	-	-	-	-	-	-
Thrombocytopenia	3	-	-	-	2	-	-	-
Hypophosphataemia	8	-	2	-	4	-	-	-

Empty cells indicate an incidence of <2%

4.2.3.1 Adverse events leading to withdrawal

Rates of dose modification due to AEs were reported as 68.2% in the regorafenib group compared with 31.1% in the placebo group.[6, 13] The rate of permanent discontinuation of the study drug due to any AE was 25% in the regorafenib group compared with 19% in the placebo group (CSR[18] Table 10-2, page 147 and Bruix *et al* 2017[6]). Any drug-related AEs led to discontinuations in 10% of patients in the regorafenib group and 4% of patients in the placebo group (CSR[18] Table 10-2, page 147 and Bruix *et al* 2017[6]). The most frequent AEs leading to discontinuation of regorafenib treatment were reported in the CS[1] (page 80) or Bruix *et al*[6] as general physical health deterioration [redacted]% in the regorafenib group compared with [redacted]% in the placebo group); increased AST (2% vs. 2%); increased blood bilirubin ([redacted]% vs. [redacted]%); hand foot skin reaction (2% vs. 0%); and ALT increase (1% vs 0%).

As reported in the principal trial publication,[6] dose interruptions or reductions due to drug-related AEs occurred in 54% of regorafenib patients and 10% of placebo patients. According to the CS[1] (page 80), dose reductions (not including interruptions) due to AEs occurred in [redacted]% of the patients in the regorafenib group and [redacted]% of the placebo group. These included hand foot skin reaction ([redacted]% in the regorafenib group compared with [redacted]% in the placebo group); diarrhoea ([redacted]% vs. [redacted]%); fatigue ([redacted]% vs. [redacted]%); and increased blood bilirubin ([redacted]% vs. [redacted]%). The most common reason for discontinuing placebo was increased AST ([redacted]% compared with [redacted]% for regorafenib).

4.2.3.2 Deaths

There were 50 deaths (13%) in the regorafenib group and 38 deaths (20%) in the placebo group. Deaths assessed as being related to the study drug were reported for seven (2%) regorafenib patients and two (1.0%) placebo patients. The seven deaths considered related to regorafenib were recorded as (one of each case): duodenal perforation, meningorrhagia, haemorrhagic shock, hepatic encephalopathy, myocardial infarction and one event for which the primary cause of death was an AE associated with clinical disease progression, for which the treating physician assessed the event as being related to the study treatment.[6]

4.2.4 *Ongoing studies*

There are currently no relevant ongoing studies of regorafenib for this indication.

4.2.5 Discussion

The company's systematic review was generally well-conducted. However, some processes could have been reported better and some relevant abstracts and additional analyses relating to the pivotal RESORCE trial should have been identified and included in the CS. This additional literature is cited, where appropriate, in this report. The review only included a single, relevant RCT: the RESORCE trial. This was an international, placebo-controlled Phase III trial evaluating the efficacy and safety of

regorafenib 160mg per day in adult patients with HCC who have previously progressed on sorafenib. RESORCE is a high quality RCT, with a low risk of selection, performance, detection, attrition and reporting bias, and with only small questions to be raised over industry involvement. The trial reported that regorafenib was significantly more effective than placebo across the primary (OS) and secondary (PFS, TTP, ORR) outcomes.

The trial found that patients on regorafenib had increased survival: median OS was reported to be 10.6 months (95% CI 9.1-12.1) in patients randomised to regorafenib compared with 7.8 months (95% CI 6.3-8.8 months) in patients randomised to placebo. The estimated HR for OS (regorafenib compared with placebo) was 0.63, 95% CI 0.50-0.79, one-sided p=0.000020. Median PFS as measured by mRECIST was significantly better for regorafenib (3.1 months, 95% CI 2.8–4.2 months) than for placebo (1.5 months, 95% CI 1.4–1.6 months): HR, 0.46, 95% CI 0.37-0.56; p<0.0001. The median TTP as measured by mRECIST was also significantly better for regorafenib (3.2 months, 95% CI 2.9–4.2) than for placebo (1.5 months, 95% CI 1.4–1.6): HR, 0.46, 95% CI 0.36-0.55; p<0.0001. The ORR, which includes both CR and PR according to mRECIST, was also significantly higher in the regorafenib group than in the placebo group (11% compared with 4%; p=0.0047). Similar findings were reported across all outcomes when using the RECIST 1.1 criteria. Subgroup analyses demonstrated consistent benefit for patients treated with regorafenib, although an additional analysis found that those who develop a new extrahepatic lesion when they progressed on sorafenib had a considerably worse survival rate than those who did not.[6]

The RESORCE trial also found that HRQoL was consistently worse on treatment than on placebo across different measures: these differences were found to be statistically significant in the case of the FACT-Hep total and the Trial Outcome Index, but did not reach clinical significance according to pre-specified thresholds.

AEs were frequent: 100% of regorafenib patients receiving the study drug experienced at least one AE (compared with 93% on placebo), and 93% of regorafenib patients experienced treatment-emergent drug-related AEs compared with 52% of placebo patients. The principal AEs were: hand foot skin reaction (53% in the regorafenib arm compared with 8% in the placebo arm); diarrhoea (41% vs 15%); fatigue (40% vs 32%); hypertension (41% vs 6%); and anorexia (31% vs 15%). AEs of Grade 3 or higher were reported for 80% of patients in the regorafenib group compared with 59% in the placebo group. Many more regorafenib patients than placebo patients also experienced Grade 3 (46% compared with 16%) and Grade 4 (4% compared with 1%) drug-related AEs. The incidence of haemorrhage events of ≥ Grade 3 was higher in the placebo group (8%) than the regorafenib group (6%), but the incidence of drug-related haemorrhage events of ≥ Grade 3 was higher in the regorafenib group (1.6%)

than the placebo group (0%). According to the CS, drug-related severe AEs were relatively low in both groups, but higher in regorafenib-treated patients compared with those receiving placebo (10% vs. 3%).

Sixty-eight percent of regorafenib patients had dose interruptions or reductions due to AEs compared with 31% of placebo patients, and dose interruptions or reductions due to drug-related AEs occurred in 54% of regorafenib patients and 10% of placebo patients. According to the CS, dose reductions (not including interruptions) due to AEs occurred in [redacted]% of the patients in the regorafenib group and [redacted]% of the placebo group. The AE profile of regorafenib in the RESORCE trial is generally similar to that of regorafenib in trials in colorectal cancer[29, 30] and there does not appear to be a statistically significant relationship between exposure and treatment-emergent AEs.[15] Deaths assessed as related to the study drug were reported for seven (2%) regorafenib patients and two (1%) placebo patients. There are no relevant ongoing studies of regorafenib.

The principal issue with the evidence concerns the limits of the trial population and how far they reflect the population seen in clinical practice in the UK. The RESORCE trial only included meaningful data on patients who were found not to be intolerant to sorafenib, who were ECOG PS 0 or 1, and who were categorised as Child-Pugh class A. The patients included in the RESORCE trial have been described as being relatively 'well'.[31, 32] A recent audit of sorafenib use in the UK[26] found that sorafenib is also used in patients who are ECOG PS 2 and Child-Pugh class B (21% and 16% of the audit population, respectively). These patients have a poorer prognosis and are more unwell. The RESORCE patients also appear to have had a substantial level of tolerance for sorafenib (at least 400mg per day for at least 20 of the last 28 days of treatment), despite rates of dose reduction/interruption and discontinuation with sorafenib being known to be relatively high.[33] The RESORCE trial patients therefore represent a particular group of adult patients with HCC who can tolerate tyrosine kinase inhibitors (TKIs) and have a relatively good prognosis.[31, 32] The licence currently includes all adult patients with HCC who have been previously treated with sorafenib. It therefore does not exclude patients who are ECOG PS 2, Child-Pugh class B, or who are intolerant to sorafenib. The CS acknowledges that there is no meaningful clinical evidence for the efficacy and safety of regorafenib in any of these groups. The sorafenib audit found that ECOG PS $\geq$2 was an independent predictor of mortality (confirming the findings of a sub-analysis of the pivotal SHARP trial[34]) and OS was substantially worse for patients who were Child-Pugh class B (4.6 months) than for those who were Child-Pugh class A (9.5 months).[26] RESORCE subgroup analyses found that patients who were PS 0 and Child-Pugh A5 experienced better efficacy than those who were PS1 and Child-Pugh A6.[6] The sorafenib audit also reported that liver dysfunction was much more common as an AE in Child-Pugh class B patients (40%) compared with Child-Pugh class A patients (18%), as was deterioration in performance status (47% vs 32%).[26] It should be noted that the number of Child-Pugh class B patients was smaller than Child-Pugh class A patients (n=43 vs n=181).[26]

Consequently, given the AE profile of regorafenib, there is a probability that patients who do not match the RESORCE trial population will experience less efficacy and more AEs (because many AEs are hepatic) than patients who match the clinical profile of the RESORCE trial population. The RESORCE trial found that HRQoL was consistently worse on treatment than on placebo across different measures and so this risk/benefit balance might be worse still for the groups without data.[31] The lack of relevant data and its implications are acknowledged in the SmPC[28]: this recognises the potential adverse impact of regorafenib on hepatic function in patients who are Child-Pugh class B and the need to monitor all AEs carefully in this group. There is therefore substantial uncertainty concerning the benefits of regorafenib in patients who do not satisfy the inclusion criteria of the RESORCE trial.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The CS does not contain an evidence synthesis of multiple studies.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

The CS does not contain an evidence synthesis of multiple studies.

4.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG did not undertake any additional analyses for the clinical effectiveness review.

4.6 Conclusions of the clinical effectiveness section

The company's systematic review was generally well conducted. The review included a single RCT: the RESORCE trial, which represents the relevant evidence. This was an international, placebo-controlled Phase III trial evaluating the efficacy and safety of regorafenib 160mg per day in adult patients with HCC who have previously progressed on sorafenib. RESORCE is a high quality RCT, with a low risk of selection, performance, detection, attrition and reporting bias. The trial reported that regorafenib was significantly more effective than placebo across the primary (OS) and secondary (PFS, TTP, ORR) outcomes. Subgroup analyses demonstrated consistent benefit for patients treated with regorafenib, but the trial also found that HRQoL was consistently worse on treatment than on placebo across different measures: these differences were found to be statistically significant in the case of the FACT-Hep total and the TOI, but did not reach clinical significance. AEs were frequent. The principal issue with the evidence concerns the limits of the trial population and how far they reflect the population seen in clinical practice in the UK. The RESORCE trial only included meaningful data on patients who were found not to be intolerant to sorafenib, who were ECOG PS 0 or 1, and who were categorised as Child-Pugh class A. The efficacy and safety of regorafenib in other adult HCC patients covered by the NICE scope and the licence, that is, those who are intolerant to sorafenib, or who are Child-Pugh class B or ECOG PS 2, is uncertain.

5 COST EFFECTIVENESS

This chapter presents a summary and critical appraisal of the methods and results of the company's review of published economic evaluations and the *de novo* health economic analysis presented within the CS.[1]

5.1 ERG comment on the company's systematic review of cost-effectiveness evidence

5.1.1 Description of company's systematic review of cost-effectiveness evidence

The company undertook an initial and simultaneous search to identify economic evaluations, cost and resource use and HRQoL of patients with advanced HCC. The following sources were searched: MEDLINE, EMBASE, EconLit, ABI/INFORM [via ProQuest], Cochrane Database of Systematic Reviews [via Wiley], Database of Abstracts of Reviews of Effects [via Wiley], and Cochrane Central Register of Controlled Trials [via Wiley], Cochrane Methodology Register [via Wiley], Health Technology Assessments Database [via Wiley] and NHS Economic Evaluation Database [via Wiley].

The company carried out supplementary searches in several international health technology assessment (HTA) agency websites (NICE, Scottish Medicines Consortium, National Centre for Pharmacoeconomics, Canadian Agency for Drugs and Technology/ pan-Canadian Oncology Drug Review, Haute Autorité de santé, Gemeinsame Bundesausschuss, Institute for Quality and Efficiency in Health Care, Dental and Pharmaceutical Benefits Board Agency).

Conference proceedings websites were searched for abstracts covering the period from 2014 to January 2017 (AACR, ASCO, Gastrointestinal Cancers Symposium, ESMO, ILCA, ESDO, EASL, ESMO World Congress on Gastrointestinal Cancer, JSMO, CSCO, and AASLD).

The company carried out separate update searches for economic evaluations and costs and resource use from July 2016 to May 2017. The HRQoL search was undertaken up to January 2017.

The ERG considers that the search was comprehensive and clearly and fully reported in CS Appendices G-I.[9]

The company's search initially identified 23 publications. Four of these studies were dismissed, leaving 19 economic evaluations. Most of the included studies related to sorafenib for HCC; none of the included studies assessed the cost-effectiveness of regorafenib for HCC.

The ERG conducted an update search in MEDLINE and EMBASE [via Ovid] on 27th July 2017. This search identified one published economic evaluation of regorafenib versus BSC for the treatment of

advanced HCC (Parikh *et al*[35]). This publication was accepted for publication in May 2017, after the company's final literature searches had been performed, hence it could not have been identified by the company's search strategy. The study by Parikh *et al*[35] compares regorafenib versus BSC for patients with unresectable HCC and Child Pugh A cirrhosis from the US health system perspective. The authors estimated that the ICER for regorafenib compared with BSC was $224,362 per QALY gained. As part of their response to the ERG's clarification questions (see company's clarification response,[8] question B17), the company stated that the study reported by Parikh *et al*[35] would have been of little value for the current appraisal as: (i) the evaluation was conducted over a restricted time horizon; (ii) no extrapolation of the data obtained in the RESORCE trial was performed thereby underestimating the survival gain associated with regorafenib, and (iii) the evaluation was performed for the US healthcare system using the US cost for regorafenib without the benefit of the UK PAS. The ERG agrees that this published study cannot adequately address the decision problem set out in the final NICE scope.[36]

5.2 Description of the company's model

5.2.1 Model scope

As part of its submission to NICE, the company submitted a fully executable health economic model programmed in Microsoft Excel®. The scope of the company's economic analysis is summarised in Table 18. The company's model assesses the cost-effectiveness of regorafenib (plus BSC) versus BSC alone for adult patients with unresectable HCC who have been previously treated with sorafenib. Incremental health gains, costs and cost-effectiveness of regorafenib are evaluated over a 15-year time horizon from the perspective of the UK NHS and Personal Social Services (PSS). All costs and health outcomes are discounted at a rate of 3.5% per annum. Unit costs are valued at 2015/16 prices.

Table 18: Summary of company's health economic model scope

Population	Adult patients with unresectable HCC who have been previously treated with sorafenib
Intervention	160mg regorafenib once daily for 3 weeks, followed by 1 week off therapy (plus BSC). Dose reductions and treatment interruptions are also included.
Comparator	BSC
Primary health economic outcome	Incremental cost per QALY gained
Perspective	NHS and PSS
Time horizon	15 years
Discount rate	3.5% per year
Price year	2015/2016

NHS – National Health Service; PSS – Personal Social Services

Population

The population considered within the company's economic analysis relates to adults with HCC who have been previously treated with sorafenib. This is broader than the population recruited into the RESORCE trial,[6] which excluded patients who discontinued sorafenib due to toxicity rather than progression (see Section 4.2) as well as those with Child-Pugh class B disease and those with an ECOG PS of 2 or more. Within the RESORCE trial, patients randomised to the regorafenib group had a mean age of 61.8 years and 87.9% of patients were male; patients randomised to the placebo group had a mean age of 61.1 years and 88.1% of patients were male. A detailed breakdown of patient characteristics is presented in Table 7.

Intervention

The intervention under consideration is regorafenib (given alongside BSC). Regorafenib is assumed to be administered orally at a dose of 160mg once daily (4 x 40mg tablets) for the first 21 days of each 28-day treatment cycle; no treatment is taken during the remaining 7 days of the cycle. The SmPC for regorafenib[28] notes that dose interruptions and/or reductions may be required and should be applied in decrements of 40mg (one tablet), with a lowest recommended daily dose of 80mg. The company's model includes such dose reductions and treatment interruptions based on an analysis of patient-level data from the RESORCE trial,[6] calculated using the mean daily dose of regorafenib prior to and after disease progression.

The SmPC for regorafenib[28] states that *"Treatment should continue as long as benefit is observed or until unacceptable toxicity occurs."* This is consistent with the RESORCE trial in which a proportion of patients who were randomised to the intervention group were allowed to continue treatment with regorafenib post-progression if the treating physician considered that the patient was still experiencing clinical benefit. As such, the company's model includes the costs of post-progression regorafenib use

in the intervention group. The ERG notes that the inclusion of these post-progression regorafenib costs is internally consistent with the experience of the RESORCE trial as the health benefits included in the model are aligned with the resources consumed to generate those benefits. The CS states that many physicians would not treat patients following disease progression,[1] therefore the scenario assessed by the model may deviate somewhat from usual clinical practice in England. The clinical advisors to the ERG did not have a consensus on whether regorafenib would be used in England for patients following progression, in line with the RESORCE trial.[6]

Comparator

The comparator included in the company's model is BSC. This is assumed to consist of concomitant medications, antibiotics, analgesics, radiation therapy for pain control (for bone metastases only), corticosteroids, transfusions, psychotherapy, growth factors, palliative surgery or any other symptomatic therapy necessary to provide BSC, except investigational anti-tumour agents, immunotherapy, antineoplastic chemotherapy or antineoplastic hormone therapy (CS,[1] page 33).

The company's base case model does not include post-progression regorafenib treatment in the comparator group as treatment switching was not permitted during the double-blind phase of the RESORCE trial.

Model versions and revisions

Three versions of the company's model were received by the ERG, two of which were submitted following the clarification process. These models include:

(1) The company's original submitted model: this was provided as part of the company's submission. This model is based on the 29th February 2016 DCO of the RESORCE trial.

(2) A revised model (submitted following clarification). This model is also based on the 29th February 2016 DCO of the RESORCE trial. This model includes additional functionality to allow for modelling of independent OS curves for each treatment group and the parametric modelling of time to treatment discontinuation used to estimate regorafenib acquisition costs. This version of the model also includes minor amendments to some of the model parameters, based on issues identified by the ERG during the clarification process (see Section 5.3).

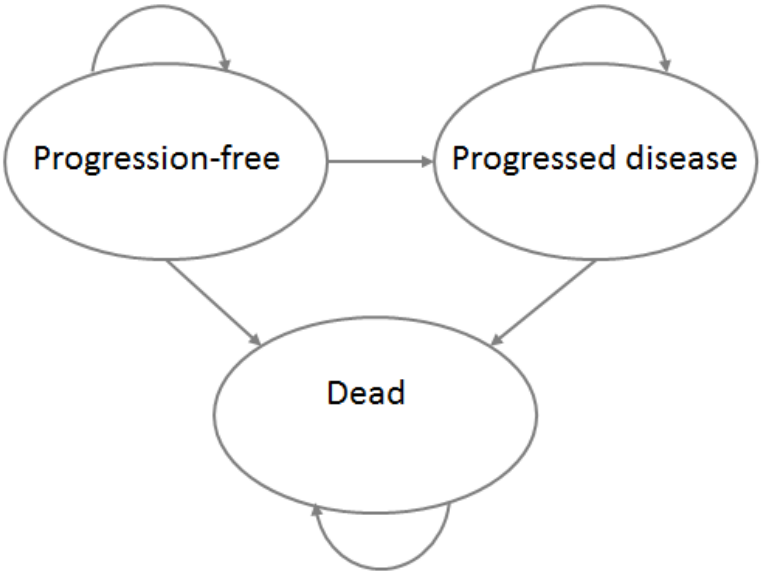
(3) A revised base case model based on the latest DCO of the RESORCE trial (23rd January 2017, submitted following clarification). This model includes updated analyses of time-to-event data; however, the model functionality is restricted to only allow for the modelling of dependent OS curves (including a treatment covariate). Following the identification of a programming error, the company submitted a corrected version of this revised base case model.

The scope of all three models is the same. Throughout this section, the model summary, results and critique relate to the original submitted model, unless otherwise stated. Details of the two models submitted following the clarification process are presented subsequently.

5.2.2 Description of the company's health economic model structure and logic

The general structure of the company's model is presented in Figure 8. The CS states that the company's model adopts a Markov approach; however, this is not an accurate description of the implemented model. Rather, the company's model adopts a partitioned survival approach based on three health states: (1) progression-free; (2) progressed disease, and (3) dead.

Figure 8: Company's model structure



The model operates as follows. Patients enter the model in the progression-free state and receive treatment with regorafenib plus BSC or BSC alone. The probability of being alive and progression-free at any time *t* is taken directly from the observed Kaplan-Meier PFS curves from the RESORCE trial (29 February 2016 DCO, prior to the permitted use of open-label regorafenib).[6] The probability of being alive at any time *t* is modelled using parametric survivor functions (log normal) including a treatment covariate (a hazard ratio [HR] derived from a Cox model) which were fitted to the individual patient-level data (IPD) from the RESORCE trial.[6] The probability of being in the post-progression state at any time *t* is calculated as the difference between the cumulative survival probabilities for OS and PFS. The model is evaluated using 28-day cycles with costs and health outcomes evaluated over a total of 195 cycles (approximately 15 years). Half-cycle correction is applied to account for the timing of events. The model includes post-progression regorafenib treatment for a proportion of patients in the intervention group, based on an analysis of IPD from the RESORCE trial.

HRQoL is principally determined by the presence/absence of disease progression. Disutilities associated with AEs are included for both groups during both the progression-free and post-progression phases. Health utilities are not adjusted by age.

The model includes costs associated with: (i) drug acquisition; (ii) hospitalisation; (iii) medical staff visits; (iv) laboratory test, and (v) radiological tests. Drug costs are adjusted according to the mean daily dose of regorafenib received pre-progression and post-progression. Health state resource use estimates were derived from a survey conducted in 2015 with three leading clinical experts in the field of oncology in the UK. As the model includes post-progression regorafenib treatment for a proportion of patients in the intervention group, costs associated with drug acquisition and associated AEs are included during both the progression-free and post-progression intervals for the intervention group.

The application of different PFS and OS curves, AE rates and treatment costs and other resource use leads to different profiles of costs and health outcomes for the two treatment groups. Incremental cost-effectiveness is calculated as the difference in costs divided by the difference in QALYs for regorafenib and BSC.

Key structural assumptions employed in company's model

The company's model employs the following key structural assumptions:

- The probability of being alive and progression-free over time for regorafenib and BSC is derived from the observed time-to-event PFS outcomes observed within the RESORCE trial.[6] Parametric curves were not fitted to these data as the cumulative PFS probabilities reach zero within the observed period of the trial for both groups.
- Within the company's base case, OS probabilities for each treatment group are modelled using log normal functions. A treatment covariate (an HR) is applied to the regorafenib group (as the baseline) to estimate OS probabilities for the BSC group.
- The company's model includes the costs associated with the use of post-progression regorafenib for a proportion of patients in the intervention group; this reflects the experience of the RESORCE trial.[6] The CS[1] (page 116) notes that the proportion of patients receiving treatment beyond recurrence in the trial is expected to be higher than would occur in clinical practice in England.
- The model includes dose reductions and treatment interruptions to manage AEs based on the experience of the RESORCE trial.[6] This assumption leads to reductions in the acquisition costs of regorafenib. This is assumed to reflect clinical practice in England.
- The model assumes that only AEs of Grade 3/4 severity are associated with impacts on costs and HRQoL. AEs occurring in ≥5% patients in either group of the RESORCE trial population

were included (anaemia, ascites, AST increase, blood bilirubin increase, fatigue, hypertension, hypophosphatemia, and palmar-plantar erythrodysaesthesia syndrome).

- The probability of receiving post-progression regorafenib and the post-progression treatment continuation rate are assumed to be independent of the time of disease progression.

5.2.3 Evidence used to inform the model parameters

Table 19 summarises the evidence sources used to inform the parameters of the company's model. The derivation of the model parameter values using these sources is described in further detail in the following sections.

Table 19: Summary of evidence sources used to inform the model parameters

Parameter type	Parameter	Source(s)
Time-to-event probabilities	PFS – regorafenib	Observed Kaplan-Meier PFS curves from RESORCE[6]
	PFS – BSC	
	OS – regorafenib	Log normal function fitted to IPD for regorafenib group in RESORCE[6]
	OS – BSC	Log normal function fitted to IPD for regorafenib group in RESORCE[6] including a treatment effect covariate (HR)
AEs	AE rate - regorafenib	Analysis of IPD from RESORCE[6]
	AE rate – BSC	Analysis of IPD from RESORCE[6]
HRQoL	Health utility – progression-free state	Tobit regression of data from RESORCE[6]
	Health utility – progressed disease state	
	Disutility – AEs	
Mean dosing	Mean daily regorafenib dose pre-progression	Analysis of IPD from RESORCE[6]
	Mean daily regorafenib dose post-progression	
Treatment continuation rates	Discontinuation probability per cycle whilst progression-free	Based on proportion of patients discontinuing regorafenib for more than one cycle prior to disease progression and median PFS from RESORCE[6]
	Probability of continuing regorafenib post-progression and duration of use	Based on proportion of patients who continued to receive regorafenib after disease progression and post-progression treatment rate in RESORCE[6]
Health state resource use	Visits, tests and hospitalisations. Separate estimates applied for: (1) Progression-free (treated with regorafenib) (2) Progression-free (treated with BSC) (3) Additional resources used at time of progression for regorafenib (4) Post-progression (treated with regorafenib) (5) Post-progression (treated with BSC)	Survey of resource use associated with sorafenib (3 clinical experts)[9]

Parameter type	Parameter	Source(s)
Unit costs	Regorafenib acquisition cost (including PAS)	Bayer[1]
	Unit costs for visits, appointments, hospitalisations, laboratory tests, radiological tests and AEs	NHS Reference Costs 2015/16,[37] Personal Social Services Research Unit (PSSRU) 2016,[38] Akhtar and Chung,[39] Cardiff and Vale Acute Chemistry Repertoire 2016/17 NHS Standards and Indicators,[1] Freedom of Information Act request[1]

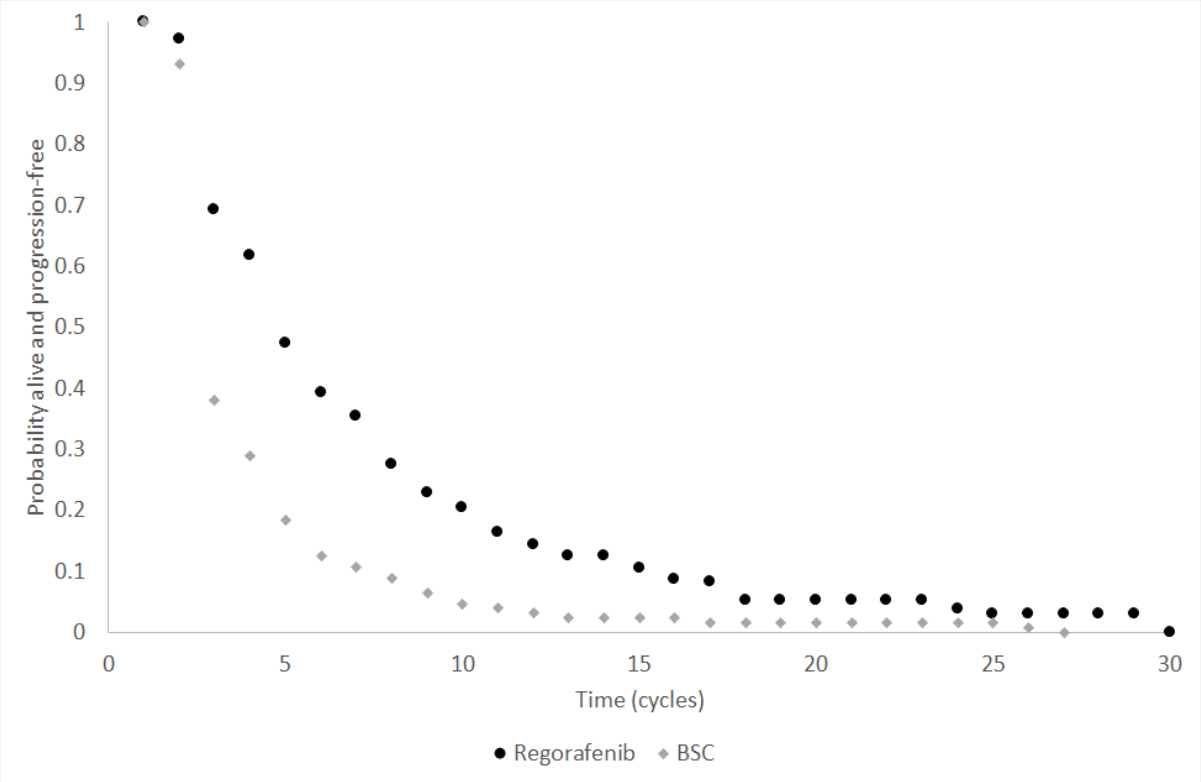
IPD – individual patient-level data; PAS – Patient Access Scheme; PSSRU – Personal Social Research Unit

Time-to-event parameters

Progression-free survival

Within the RESORCE trial, PFS was defined as the time from the date of randomisation to the date of disease progression (radiological or clinical) or death due to any cause (whichever occurred first) (CS,[1] Table 14). Within the company's model, PFS was based on mRECIST criteria. The CS[1] notes that mRECIST *"includes amendments developed for the SHARP trial (27) that require cytopathological confirmation of malignancy to classify pleural effusion or ascites as progression, and that apply more stringent criteria to define progression due to lymph node involvement at the hepatic hilum or new intrahepatic sites (28). It also considers complete tumour necrosis on dynamic imaging studies"* (CS,[1] page 43). Given that the Kaplan-Meier curves for PFS within RESORCE show the complete pattern of PFS for both the intervention and control groups (the cumulative PFS probabilities drop to zero in both groups within the observed period of the trial), the company's model uses these observed time-to-event data directly: parametric survival curves were not fitted to available data. The observed PFS probabilities for each model cycle are presented in Figure 9. Based on these data, the model estimates undiscounted mean PFS durations of 0.47 years for the regorafenib group and 0.23 years for the BSC group.

Figure 9: mRECIST PFS probabilities used in the company's model (derived from company's model)

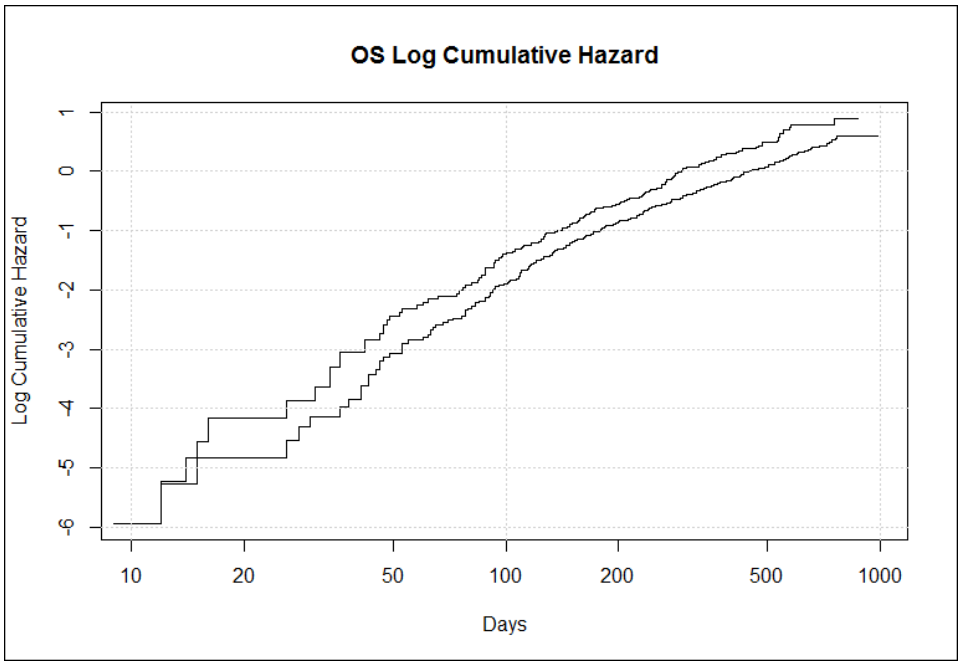


Overall survival

OS is defined as time from the date of randomisation until the date of death due to any cause. Exponential, Weibull, Gompertz, log normal, log logistic and gamma functions were fitted to the available OS time-to-event data. The company explored the use of both independent survival models (fitted separately to each treatment group) and joint models (including a treatment covariate for placebo applied to the regorafenib group as a baseline). Based on an inspection of the log cumulative hazard plots, the company concluded that although the traces for regorafenib and BSC cross at around the 15-day time point, these appear otherwise parallel (see Figure 10). In addition, the company conducted a Grambsch and Therneau test between the Schoenfeld residuals and the log of time: this analysis produced a non-significant *p*-value (p=0.331) which suggests that the proportional hazards assumption is not violated. On the basis of these two pieces of information, the CS argues that the proportional hazards assumption is plausible and the company's base case analysis is based on the jointly fitted model (dependent OS curves). It should be noted however that the treatment effect parameter derived from the jointly fitted models is not used in the company's base case model; instead, an HR derived from a Cox model is used for all parametric model types. The ERG notes that this is inappropriate for accelerated failure time (AFT) models as the treatment effect covariate reflects a constant acceleration factor rather than an HR (see Section 5.3).

Model discrimination was undertaken through examination of goodness-of-fit statistics (the Akaike Information Criterion [AIC] and the Bayesian Information Criterion [BIC]), visual inspection and what is referred to within the CS as "clinical validity." All of these analyses relate only to the jointly fitted model which includes a treatment effect covariate: the CS does not include goodness-of-fit statistics or survival plots for the independently fitted models, although these were provided in response to a request for clarification (presented in Section 5.3). The ERG also notes that the company's exploration of clinical validity does not make reference to the use of subjective clinical judgements about the plausibility of the extrapolation beyond the observed period within the RESORCE trial.[6]

Figure 10: Log cumulative hazard plot for OS (reproduced from CS Figure 14)



Top curve – BSC; bottom curve – regorafenib

The observed and modelled OS predictions parametric OS functions are presented in Figure 11. AIC and BIC statistics for the jointly fitted models are shown in Table 20.

Table 20: Overall survival – AIC and BIC statistics from jointly fitted parametric models (adapted from CS Table 35)

Survivor function	AIC	BIC
Log normal	**5197.513**	**5210.565**
Log logistic	5199.734	5212.787
Gamma	5211.014	5224.067
Weibull	5218.877	5231.929
Gompertz	5238.261	5251.314
Exponential	5239.994	5248.696

Lowest values highlighted in bold

With respect to the statistical goodness-of-fit of the OS models, both the AIC and BIC statistics were lowest for the log normal function (AIC=5197.513, BIC=5210.565). The AIC and BIC were markedly higher for all other models except the log logistic function (AIC=5199.734, BIC=5212.787). The ERG notes that this relates only to the fit of the model to the observed data, which in isolation, represents an insufficient basis for selecting a preferred model.

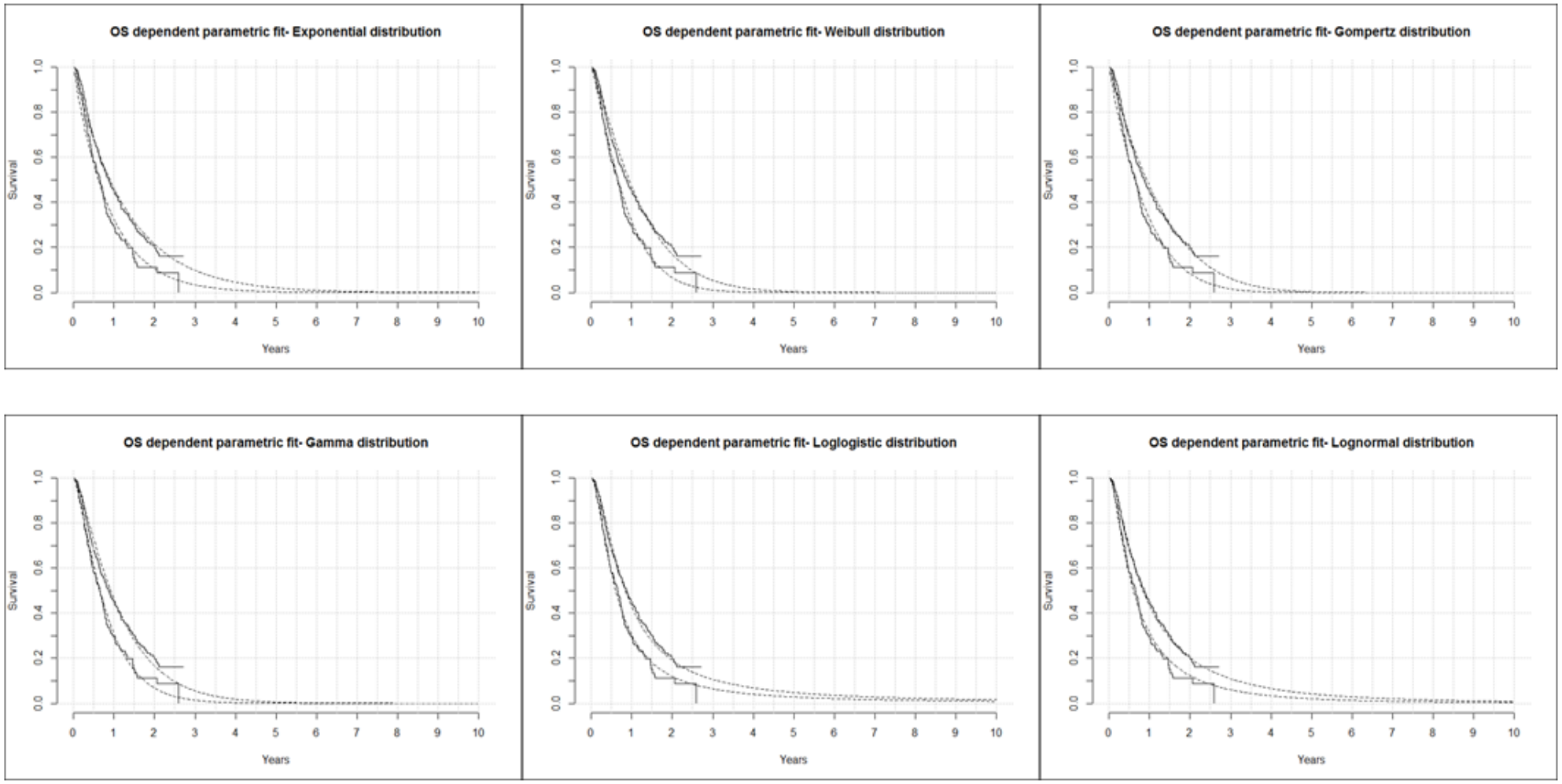
As discussed in the CS, all of the parametric survivor functions considered appear to provide a reasonably good visual fit to the observed OS data, and the log normal and log logistic functions have longer tails than the other candidate parametric functions (see Figure 11).

Within the CS, the company isolates three parametric functions as being potentially clinically valid: the exponential, the log normal and the log logistic models. With respect to the clinical validity of the parametric curve-fitting, the CS notes the following:

- At the 35-cycle timepoint (the last time point for which observed data were available), the log logistic and log normal curves provided the closest fit to the observed Kaplan-Meier curves for both the regorafenib and placebo groups. The exponential curves provide a good approximation for the regorafenib group, but not for the placebo group.
- At the 5-year time point, the log logistic and log normal functions predict a small probability of survival in the BSC arm (log logistic OS probability = 0.03; log normal OS probability = 0.02) and a greater (but small) probability of survival in the regorafenib arm (log logistic OS probability = 0.05; log normal OS probability = 0.04).
- At the 10-year time point, the exponential function predicts approximately zero survival in both the regorafenib and BSC groups, the log logistic function predicts a small survival probability in both groups (OS probability = 0.02 and 0.01 for regorafenib and BSC, respectively) and the log normal function predicts a very small survival probability for regorafenib (OS probability = 0.01 and 0.00 for regorafenib and BSC, respectively).

On the basis of the above information, the log normal model was selected for inclusion in the company's base case analysis. The CS includes sensitivity analyses using each of the alternative parametric OS functions.

Figure 11: Overall survival – parametric curve fits from jointly fitted parametric models (reproduced from CS Figure 15)



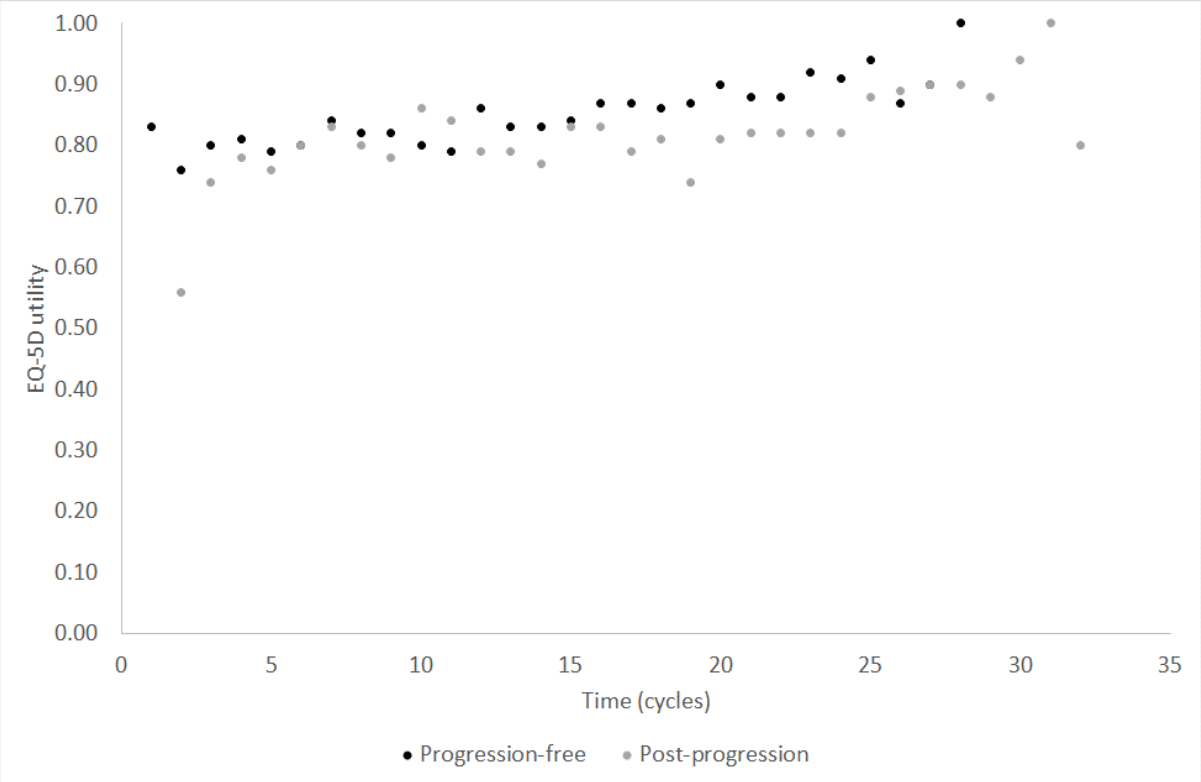
The dotted curves are the fitted extrapolations, and the solid curves are the raw trial data. The upper curves represent the regorafenib arm. The lower curves represent the placebo arm

Health-related quality of life

The company undertook a systematic review of studies reporting HRQoL in HCC (see CS Appendix H[9]). The company considered only those studies which reported HRQoL for HCC patients for both progression-free and post-progressed states (presumably for the sake of consistency, although this is not explicitly stated within the CS). The company's HRQoL review identified only one relevant published study: the placebo-controlled SHARP trial of sorafenib for HCC.[21] The CS notes that the utilities derived from SHARP lack face validity as the value for the post-progression state is higher than that for the progression-free state (progression-free utility=0.6885; post-progression utility=0.7111). These data are not used in the company's base case analysis. Instead, the health utilities employed in the company's model were derived from EQ-5D data collected within the RESORCE trial.[6]

Within the RESORCE trial, EQ-5D questionnaires were self-administered by patients at the start of each treatment visit (the first day of each treatment cycle) whilst the patient was receiving blinded treatment and before they saw the investigator or any study-related procedures were performed. An additional EQ-5D assessment was completed during the "end of treatment" visit.[1] Mean EQ-5D utilities for the pooled treatment population are presented in Figure 12. As shown in the figure, the mean EQ-5D utility at most of the assessment points remains generally high (progression-free utility range = 0.76 to 1.0; post-progression utility range = 0.56 to 0.90).

Figure 12: EQ-5D utility by treatment cycle (both groups pooled, excludes end of treatment visit assessment, point estimates only)



In response to a request for clarification from the ERG (see company's clarification response,[8] question B10), the company provided additional data showing the EQ-5D completion rates for the progression-free and post-progression health states at each assessment visit (see Table 21). The ERG notes that the EQ-5D completion rates were consistently high for the progression-free state, but were subject to considerable attrition following disease post-progression.

Table 21: EQ-5D questionnaire completion rates over time

Cycle	Pre-progression			Post-progression		
	Completed EQ-5D	Alive	Percentage	Completed EQ-5D	Alive	Percentage
1	531	573	93%		0	
2	489	518	94%	4	44	9%
3	283	297	95%	64	224	29%
4	228	255	89%	50	226	22%
5	168	181	93%	60	250	24%
6	123	136	90%	61	248	25%
7	98	118	83%	51	215	24%
8	78	85	92%	48	219	22%
9	65	66	98%	39	211	18%
10	53	56	95%	38	191	20%
11	45	46	98%	37	177	21%
12	33	36	92%	38	159	24%
13	31	31	100%	33	146	23%
14	28	29	97%	24	128	19%
15	20	20	100%	29	116	25%
16	17	20	85%	23	97	24%
17	13	16	81%	20	88	23%
18	10	16	63%	20	79	25%
19	10	10	100%	19	77	25%
20	9	10	90%	16	72	22%
21	9	10	90%	12	54	22%
22	9	10	90%	11	43	26%
23	9	9	100%	11	40	28%
24	6	6	100%	13	35	37%
25	5	6	83%	10	29	34%
26	5	5	100%	7	23	30%
27	3	5	60%	7	20	35%
28	2	3	67%	7	18	39%
29	1	2	50%	6	14	43%
30				5	13	38%
31				3	11	27%
32				1	9	11%

Eight separate regression models were fitted to the available EQ-5D data including combinations of three covariates: (i) treatment group; (ii) progression status, and (iii) AEs. These eight models were each evaluated using: (a) Ordinary Least Squares (OLS) regression; (b) Tobit regression with repeated measurements, and (c) a mixed model for repeated measurements. This resulted in a total of 24 models

being considered. The choice of covariates for inclusion in the final model was explored using a stepwise selection approach. Across the first seven regression models, allocated treatment was not statistically significant (p>0.05), progression status was always significant and AEs were significant in most models. Goodness-of-fit was explored using the adjusted R-squared for the OLS models and using the AIC and BIC for the Tobit and mixed models. The company's preferred analysis was the Tobit model including covariates for progression status and AEs ("Model 8" in CS[1] Table 42). The health utility values applied in the company's model are summarised in Table 22.

Table 22: Health utilities applied in the company's model

Health state / event	Mean utility	SE
Progression-free	0.811	0.00
Progressed disease	0.763	0.01
AEs (disutility)	-0.014	0.01

Regorafenib treatment dose during progression-free phase

The cost of regorafenib treatment during the progression-free and post-progression phases was estimated according to the mean daily dose within the RESORCE trial,[6] based on an analysis of the patient-level data from the trial (see Table 23).

Table 23: Mean daily dose of regorafenib assumed in the company's model

Progression status	Mean daily dose (mg/day)
Progression-free	[redacted]
Post-progression	[redacted]

Regorafenib discontinuation during progression-free phase

During the progression-free phase, [redacted] of the patients are assumed to discontinue regorafenib treatment during each model cycle, based on the proportion of patients who discontinued treatment for more than one cycle prior to disease progression ([redacted]) and the median PFS duration (3.1 months).

Regorafenib continuation during post-progression phase

[redacted] percent of patients who progress are assumed to continue regorafenib treatment after progression, based on the experience of the RESORCE trial.[6] This is applied in the model based on the proportion of progressed patients who received post-progression regorafenib and the sumproduct of the proportion of patients who are new progressors during each cycle and the post-progression regorafenib continuation rate (i.e. the proportion of patients who received 1, 2, 3 etc. additional cycles of regorafenib subsequent to disease progression, see Table 24). This "cycle-cohort simulation"[8] approach assumes that the probability of receiving post-progression treatment and the post-progression treatment continuation rate are independent of the time of disease progression. The ERG notes that this aspect of

the model is not well explained in the CS and the approach taken is overly complex and makes unnecessary assumptions where observed data could have been used instead (see Section 5.3).

Table 24: Post-progression treatment rate (applied to those progressing patients who receive post-progression regorafenib treatment, reproduced from CS Table 40)

Cycle after progressing	Proportion of patients receiving *n* cycles post-progression	
	Regorafenib	BSC*
1	[redacted]	[redacted]
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		

** Post-progression regorafenib use is included within the BSC group in order to estimate impacts of AEs on costs and HRQoL. The ERG does not consider this to be appropriate, however its impact on the ICER is negligible*

Adverse event frequency

The company's model includes the following AEs: anaemia; ascites; AST increase; blood bilirubin increase; fatigue; hypertension; hypophosphataemia and palmar-plantar erythrodysaesthesia syndrome. The model assumes that AEs may be experienced during any cycle and that these impact upon both costs and HRQoL. AE rates were derived from an analysis of IPD from the RESORCE trial.[6] The model assumes per cycle probabilities of AEs of 5.55% for regorafenib and 5.06% for BSC. The distribution of AEs within each group are summarised in Table 25.

Table 25: Proportions of patients experiencing adverse events and associated costs

AE type	AE unit cost	Proportion AEs of type - regorafenib	Proportion AEs of type - BSC	AE cost regorafenib	AE cost BSC
Anaemia	£1,283.67 [a]	0.07	0.11	£84.72	£145.05
Ascites	£1,667.00 [b]	0.05	0.17	£83.35	£275.06
AST increase	£1,667.00 [b]	0.16	0.32	£261.72	£533.44
Blood bilirubin increase	£1,667.00 [b]	0.11	0.19	£178.37	£308.40
Fatigue	£1,667.00 [b]	0.08	0.05	£125.03	£85.02
Hypertension	£729.87 [c]	0.25	0.10	£178.82	£75.18
Hypophosphatemia	£1,261.96 [d]	0.14	0.04	£171.63	£51.74
Palmar-plantar erythrodysaesthesia syndrome	£873.37 [e]	0.16	0.02	£142.36	£18.34
Total	-	1.00	1.00	**£1,225.99**	**£1,492.22**

[a] Unweighted mean of total NHS Reference Costs 2015/16[37] for HRG codes SA04G to SA04L (iron deficiency anaemia with CC scores 0 to 14+).

[b] Unweighted mean of total NHS Reference Costs 2015/16[37] for HRG codes GC12G to GC12K (malignant, hepatobiliary or pancreatic disorders, without interventions, with cc score 0 to 6+).

[c] Costed using the total NHS Reference Costs 2015/16[37] for HRG codes EB04Z (hypertension).

[d] Unweighted mean of total NHS Reference Costs 2015/16[37] for HRG codes KC05J to KC05N (fluid or electrolyte disorders with interventions with cc score 0 to 10+).

[e] Costed using the total NHS Reference Costs 2015/16[37] for HRG codes XD57Z (skin conditions, drugs band 1).

Resource use and costs

The company's model includes resource costs associated with: (i) drug acquisition for regorafenib; (ii) health state resource use, and (iii) the management of AEs.

Drug acquisition

Drug acquisition costs for regorafenib were provided by the company. The company has a PAS in place for regorafenib resulting in a price of [redacted] for a packet of 84 x 40mg tablets. The recommended dose of regorafenib is 160mg for each of the first 21 days of a 28-day treatment. Within the company's model, dose reductions and treatment interruptions result in a mean daily dose of [redacted]mg per day during the progression-free phase and [redacted]mg per day during the post-progression phase. Thus, the cost of regorafenib per 28-day treatment cycle including dose reductions and treatment interruptions is [redacted] for patients in the progression-free health state and [redacted] for patients in the post-progression health state. The model assumes that BSC is not associated with any additional drug costs; the ERG notes that this may favour the intervention group as BSC is included in both groups and survival is extended for patients receiving regorafenib. Given that regorafenib is administered orally in tablet form, no administration costs were included in the company's model.

Health state resource use

Health state resource use estimates were based on a physician survey of three leading experts in HCC (conducted in 2015) carried out for sorafenib and assumes that resource use for patients receiving regorafenib is identical to that for patients receiving sorafenib (see Table 26). This assumption was not

raised as a major concern by the clinical advisors to the ERG, but the robustness of the survey was questioned. The full survey is provided in CS Appendix O.[9]

Table 26: Resource use for patients receiving regorafenib or BSC in both the progression-free and post-progression states

Description	Unit cost	Progression-free		Post progression	
		Regorafenib*	BSC	Regorafenib*	BSC
Hospitalisation					
General ward	£801	0.07	0.25	0.08	0.25
Duration of stay (days)	-	5.83	7.00	5.25	7.00
Cost of hospitalisation [1]	-	£4,670	£5,607	£4,205	£5,607
A&E admission	£138	0.37	0.25	0.08	0.25
Hospital outpatient appointments					
Oncologist	£163	1.07	0.75	1.00	0.75
Hepatologist	£253	0.33	0.00	0.00	0.00
Gastroenterologist	£132	0.00	0.00	0.00	0.00
Clinical nurse specialist	£130	0.67	0.50	0.50	0.50
Palliative care team	£131	0.00	2.17	0.00	0.00
Macmillan nurse	£73	0.00	0.00	0.00	0.00
Follow up visits					
GP visit	£36	0.00	0.00	0.00	0.00
Nurse visit	£36	0.00	0.00	0.00	0.00
Specialist visit	£151	0.84	0.84	0.50	0.84
Tests					
Alpha fetoprotein	£3.03	1.00	0.84	1.84[2]	0.84
Liver function	£2.78	1.00	0.84	1.00[2]	0.84
Biochemistry	£1.34	1.00	0.84	1.84[2]	0.84
Complete blood count	£2.65	1.00	0.84	1.84[2]	0.84
International normalised ratio	£3.43	0.71	0.34	0.67[3]	0.34
Radiological tests					
CT scan of abdomen	£122	0.39	0.17	0.84[3]	0.17
MRI of abdomen	£238	0.00	0.00	0.00	0.00

** Estimates elicited for sorafenib assumed to apply identically to regorafenib; costs of radiology and endoscopy not included in original submitted model but later included in model received post-clarification*

[1] Calculated multiplying the estimated length of stay by the estimated cost of a bed day on a general ward (£801)

[2] 1.00 at progression

[3] 0.67 at progression

Unit costs associated with the majority of resource items included in the company's model were taken from the NHS Reference Costs 2015/2016.[37] Other cost sources included: the Personal Social and Services Research Unit (PSSRU, Curtis and Burns[38]), Akhtar & Chung[39] and other NHS sources (bibliographic details not provided in the CS[1]). Of particular note, the estimated cost of a bed day in a general ward (£801 per day) was obtained from a response to a Freedom of Information Act request;[1] this is discussed in further detail in Section 5.3. Unit costs associated with AEs are summarised in Table 25.

5.2.4 Model evaluation methods

The CS presents the results of the economic analysis in terms of the incremental cost per QALY gained for regorafenib versus BSC. The base case results are presented deterministically based on point estimates of parameters. The CS also includes the results of probabilistic sensitivity analysis (PSA), deterministic sensitivity analyses (DSA) and scenario analyses. The results of the PSA are presented in the form of a cost-effectiveness plane and cost-effectiveness acceptability curves (CEACs), based on 1,000 Monte Carlo simulations. The results of the DSA are presented in tabular form with an additional tornado diagram which is limited to the ten most influential model parameters. The distributions applied in the company's PSA are summarised in Table 27.

Table 27: Distributions applied in company's probabilistic sensitivity analyses

Parameter type	Distribution	ERG comment
PFS	Fixed	These parameters are uncertain and should be included in the PSA
OS	Fixed regorafenib baseline, HR for BSC versus regorafenib sampled from normal distribution on log scale	The baseline OS curve is uncertain and should be included in the PSA
AEs	Beta	-
HRQoL	Beta	-
Mean dosing	Gamma	-
Treatment continuation rates	Fixed	These parameters are uncertain and should be included in the PSA
Health state resource use	Gamma	-
Unit costs	Fixed	These parameters are uncertain and should be included in the PSA

5.2.5 Company's model results

Table 28 presents the central estimates of cost-effectiveness derived from the company's original submitted model. Based on the probabilistic version of the model (assuming the log normal function for OS), regorafenib is expected to generate an additional 0.37 QALYs at an additional cost of £12,311 per patient: the corresponding ICER for regorafenib versus BSC is £33,335 per QALY gained. The deterministic version of the model produces a very similar ICER of £33,437 per QALY gained for regorafenib versus BSC.

Table 28: Company's central estimates of cost-effectiveness – regorafenib versus BSC

Probabilistic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. costs**	**ICER (per QALY gained)**
Regorafenib	1.045	[redacted]	0.369	£12,311	**£33,335**
BSC	0.676	[redacted]	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. costs**	**ICER (per QALY gained)**
Regorafenib	1.044	[redacted]	0.367	£12,262	**£33,437**
BSC	0.677	[redacted]	-	-	-

Figure 13 presents the CEAC for regorafenib versus BSC. Assuming a willingness-to-pay (WTP) threshold (λ) of £30,000 per QALY gained, the company's model estimates that the probability that regorafenib produces more net benefit than BSC is 0.21. Assuming a WTP threshold of £50,000 per QALY gained, the company's model estimates that the probability that regorafenib produces more net benefit than BSC is 1.0.

Figure 13: Cost-effectiveness acceptability curves – regorafenib versus BSC (reproduced from CS Figure 17)

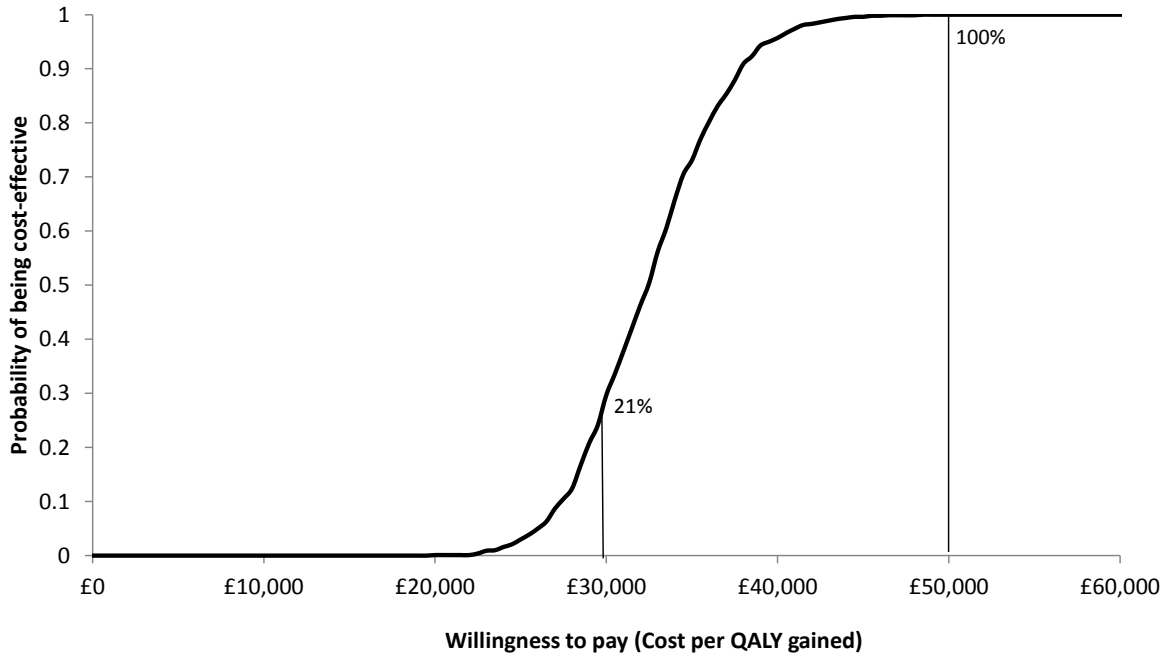
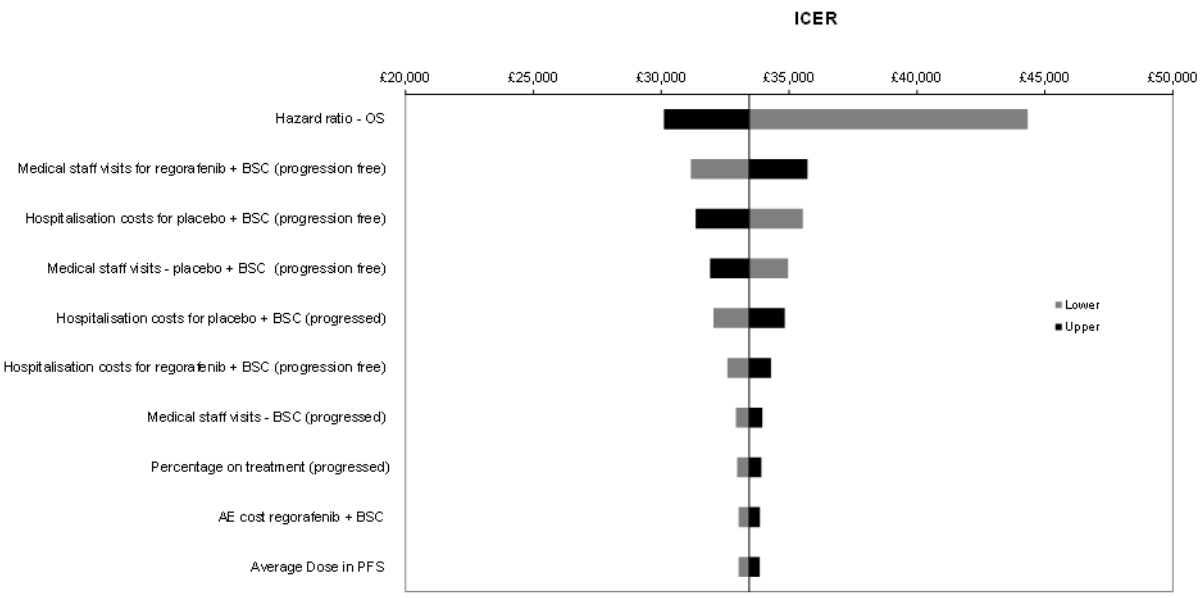


Figure 14 presents the results of the company's DSAs as a tornado diagram (mean parameters varied by +/-30%). Based on this analysis, the most influential model parameters appear to be the HR for OS and assumptions about health state resource use in both treatment groups. Across all analyses, the ICER for regorafenib versus BSC remains lower than £50,000 per QALY gained. It should be noted that the ERG was unable to replicate these analyses using the company's submitted model.

Figure 14: Results of company's deterministic sensitivity analyses – regorafenib versus BSC (reproduced from CS Figure 18*)



** The ERG was unable to replicate the tornado diagram using the submitted model*

Table 29 presents the results of the company's scenario analyses. These analyses indicate that the use of alternative parametric functions to model PFS probabilities has little impact upon the ICER for regorafenib versus BSC. The use of the generalised gamma model for OS produced a notable increase in the ICER for regorafenib plus BSC versus BSC (ICER=£39,466 per QALY gained); the use of all other parametric functions resulted in lower ICERs compared with the base case. The use of health utilities derived from the SHARP trial of sorafenib[21] increased the ICER for regorafenib versus BSC by around £4,000 per QALY gained, however both the company and the ERG have concerns regarding the face validity of these utility estimates. Doubling the disutility associated with disease progression did not have a marked impact upon the model results. The exclusion of dose reductions and interruptions (i.e. assuming a constant fixed dose of 160mg regorafenib per day) increased the ICER for regorafenib versus BSC to £41,206 per QALY gained. The exclusion or restriction of post-progression regorafenib treatment led to decreases in the ICER for regorafenib. The use of shorter time horizons increased the ICER for regorafenib. The ICER for regorafenib versus BSC remained lower than £50,000 per QALY gained across all of the company's scenario analyses.

Table 29: Company's scenario analyses – regorafenib versus BSC (adapted from CS Table 58)

	Inc. costs	Inc. QALYs	ICER (per QALY gained)
Base case	**£12,262**	**0.367**	**£33,437**
PFS – alternative parametric functions			
Log normal	£11,796	0.365	£32,302
Log logistic	£11,915	0.366	£32,571
Weibull	£12,007	0.366	£32,842
Exponential	£12,257	0.367	£33,410
Generalised gamma	£11,842	0.365	£32,456
Gompertz	£12,414	0.368	£33,775
OS – alternative parametric functions			
Log logistic	£12,755	0.395	£32,379
Weibull	£5,747	0.223	£25,726
Exponential	£7,885	0.301	£26,212
Generalised gamma	£9,692	0.246	£39,466
Gompertz	£6,768	0.245	£27,587
Utilities			
Sorafenib utility values (pre-progression = 0.6885; post-progression = 0.7111)	£12,262	0.327	£37,554
Progression disutility doubled	£12,262	0.355	£34,524
Daily average dose of regorafenib			
160mg i.e. no dose reductions or treatment interruptions	£15,111	0.367	£41,206
Post-progression treatment			
None	£10,913	0.367	£29,731*
Maximum of 3 cycles	£11,949	0.367	£32,582
Time horizon			
3 years	£9,647	0.238	£40,555
5 years	£11,004	0.305	£36,112
10 years	£12,029	0.355	£33,862

** The ERG was unable to replicate this ICER. Applying zero duration to the post-progression treatment rates gives a higher ICER of £32,194 per QALY gained*

5.3 Critical appraisal of the company's health economic analysis

This section presents a critical appraisal of the health economic analysis presented within the CS.[1] Section 5.3.1 details the methods used by the ERG to interrogate and critically appraise the company's submitted health economic analysis. Section 5.3.2 summarises the extent to which the company's analysis adheres to the NICE Reference Case.[40] Section 5.3.3 summarises the ERG's verification of the company's implemented model and highlights inconsistencies between the model, the CS,[1] and the sources used to inform the model parameter values. Section 5.3.4 presents a detailed critique of the main issues and concerns underlying the company's analysis.

5.3.1 Methods for reviewing the company's economic evaluation and health economic model

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic evaluation and the underlying health economic model upon which this was based. These included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists[41, 42] to critically appraise the company's model and analysis.
- Scrutiny of the company's model by health economic modellers and discussion of issues identified amongst the members of the ERG.
- Double-programming of the deterministic version of the company's model to fully assess the logic of the company's model structure, to draw out any unwritten assumptions and to identify any apparent errors in the implementation of the model.
- Examination of the correspondence between the description of the model reported within the CS[1] and the company's executable model.
- Replication of the base case results, PSA, DSA and scenario analyses presented within the CS.[1]
- Where possible, checking of parameter values used in the company's model against their original data sources.
- The use of expert clinical input to judge the credibility of the company's economic evaluation and the assumptions underpinning the model.

5.3.2 Adherence of the company's model to the NICE Reference Case

The company's economic evaluation is generally in line with the NICE Reference Case.[40] As discussed in Section 4, the main uncertainty regarding the scope of the company's economic analysis relates to those groups of patients who are included in the marketing authorisation for regorafenib who were excluded from the RESORCE trial.

Table 30: Adherence of the company's model to the NICE Reference Case

Element	Reference case	ERG comments
Defining the decision problem	The scope developed by NICE	The company's model is generally in line with the final NICE scope.[36] However, the ERG notes that the population included in the company's economic analysis relates to patients who have been previously treated with sorafenib, whilst the RESORCE trial[6] which is used to populate the model parameters relates to patients who have progressed on sorafenib. This study specifically excluded patients who discontinued treatment with sorafenib due to toxicity as well as those with Child-Pugh class B disease and those with an ECOG PS of 2 or more.
Comparator(s)	As listed in the scope developed by NICE	The company's choice of comparator is appropriate. BSC the only comparator listed within the final NICE scope.[36]
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health gains for patients are modelled in terms of QALYs gained.
Perspective on costs	NHS and PSS	The company's economic analysis adopts an NHS and PSS perspective.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The company's economic evaluation takes the form of a cost-utility analysis. The results of the analysis are presented in terms of the incremental cost per QALY gained for regorafenib (plus BSC) versus BSC alone.
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. Scenario analyses are also presented for alternative time horizons of 3, 5 and 10 years. The model also includes the functionality to assess the cost-effectiveness of regorafenib over a longer time horizon of approximately 18.3 years, although the impact on the company's ICER is negligible.
Synthesis of evidence on health effects	Based on systematic review	Health outcomes are based on those reported within the RESORCE trial.[6]
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	HRQoL estimates were derived from regression analyses of patient-reported EQ-5D data collected within the RESORCE trial.[6] EQ-5D responses were transformed to preference-based index utilities using the UK tariff.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	

Element	Reference case	ERG comments
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting is applied to estimated QALY gains. The company makes the case that regorafenib should be considered as a life extending treatment given at the end of life (see CS,[1] pages 93-94).
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	Resource components included in the company's model reflect those relevant to the NHS and PSS. Unit costs were valued at 2015/16 prices.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	All costs and QALYs are discounted at a rate of 3.5%.

5.3.3 Model verification and correspondence between the model, the CS and parameter sources

Double-programming of the deterministic version of the company's model

The ERG rebuilt the deterministic version of the company's base case model in order to verify its implementation. Table 31 presents a breakdown of the health outcomes and costs generated using the company's model and the ERG's rebuilt model.

Table 31: Comparison of company's base case model and ERG's rebuilt model results (undiscounted unless otherwise stated)

	ERG rebuilt model		Company's model	
Outcome	**Regorafenib**	**BSC**	**Regorafenib**	**BSC**
LYGs	1.42	0.90	1.42	0.90
LYGs (discounted)	1.34	0.87	1.34	0.87
QALYs	**1.10**	**0.70**	**1.11**	**0.70**
QALYs (discounted)	1.04	0.67	1.04	0.68
Drug costs – progression-free	[redacted]	[redacted]	[redacted]	[redacted]
Drug costs – post-progression	[redacted]	[redacted]	[redacted]	[redacted]
AE costs	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation costs	[redacted]	[redacted]	[redacted]	[redacted]
Medical visits costs	[redacted]	[redacted]	[redacted]	[redacted]
Lab tests costs	[redacted]	[redacted]	[redacted]	[redacted]
Radiological tests costs	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]
Total costs (discounted)	[redacted]	[redacted]	[redacted]	[redacted]

LYG – life year gained

As shown in Table 31, the ERG was able to produce very similar estimates of health gains to those estimated by the company. With respect to the modelled costs, the ERG identified some anomalies:

(i) The ERG identified a programming error whereby the company's formulae to estimate utility decrements due to AEs in the BSC group erroneously refers to the AE probability for patients

receiving regorafenib. This issue was rectified in the company's models submitted following the clarification process.

(ii) The model assumes that there are 13 x 28-day cycles per year. A more appropriate number of cycles is 13.044 (calculated as 365.25/28). This issue was also rectified in the company's revised models submitted following the clarification process.

(iii) Regorafenib treatment costs are calculated using the half-cycle corrected trace of health state occupancy. This assumes that patients who progress or die only receive a half-cycle worth of regorafenib rather than the full cycle's worth of treatment they would be prescribed. The acquisition costs estimated by the ERG are higher than the company's costs as they are based on the health state populations at the beginning of the model cycle.

(iv) The ERG identified some minor issues regarding the calculation of health state hospitalisation costs and medical visit costs. The main discrepancy is a consequence of two sets of programming errors in the company's model. With respect to hospitalisations, the company's model inappropriately applies a zero hospitalisation cost to patients in the regorafenib group who are progression-free but have discontinued treatment (Model worksheet "Model_cost", cells AC284:AC523): the correct hospitalisation cost that should have been applied is £848 (Model worksheet "Live", cell E68 not G68). A similar issue applies to medical visits whereby a zero cost is applied patients in the regorafenib group who are progression-free but have discontinued treatment (Model worksheet "Model_cost", cells AC534:AC773): the correct medical visit cost that should have been applied is £598 (Model worksheet "Live", cell E69 not G69). Even when these errors are corrected, there remains further discrepancy between the costs estimated by the ERG and the company: these discrepancies have the propensity to increase the ICER for regorafenib versus BSC by around £1,000.

(v) The company's model does not include half-cycle correction for the costing of AEs, thereby increasing their costs relative to those estimated by the ERG.

(vi) The AE costs for the BSC group include AEs experienced by a proportion of patients who are assumed to receive regorafenib post-progression. Given that the model does not include the acquisition costs for BSC patients switching to regorafenib, the reasons for the inclusion of these regorafenib-related AE costs are unclear.

In addition, the ERG was unable to reproduce two sets of results reported in the CS using the company's submitted model:

(i) The company's reported DSA which excludes post-progression treatment costs (CS reported ICER = £29,731 per QALY gained [see Table 29]; ERG estimated ICER = £32,194 per QALY gained).

(ii) The company's reported tornado diagram (see Figure 7). The results generated from the tornado diagram in the executable model are very different to those reported in the CS. It is unclear how the reported results were generated using the submitted model.

A further programming error was identified whereby if the Weibull OS model was selected in the company's model, the PFS trace dropped to zero in the second cycle and every cycle thereafter; this affects the company's sensitivity analyses, however the base case ICER remains unaffected.

Notwithstanding these issues and other concerns identified within the critical appraisal (see Section 5.3.4), the ERG is broadly satisfied that the company's model has been implemented as described in the CS.

Correspondence between the model inputs and their original sources

The ERG is satisfied that the inputs applied in the model reflect those described in the CS.[1] However, the ERG did not have access to the raw data used to inform the statistical time-to-event models or the Tobit EQ-5D regression model and therefore cannot verify the accuracy of their implementation.

5.3.4 Main issues identified within the critical appraisal

Box 1 summarises the main issues identified within the ERG's critical appraisal of the company's economic analysis. These issues are discussed in further detail in the subsequent sections.

Box 1: Summary of main issues identified within the company's model

(1) Inappropriate use of a hazard ratio to model treatment effects for OS
(2) Limited consideration of clinical plausibility of extrapolated OS curves
(3) Concerns regarding the modelling of time to treatment discontinuation to estimate regorafenib acquisition costs
(4) Inclusion of potentially unrealistic cost savings due to dose reductions/interruptions
(5) Concerns regarding expert clinician survey to inform health state resource use
(6) Likely overestimation of the cost of a general ward bed day
(7) Use of potentially inappropriate NHS Reference Costs
(8) Questionable reliability of post-progression utility estimate
(9) Inadequate consideration of uncertainty

(1) Inappropriate use of a hazard ratio to model treatment effects for OS

The company elected to use jointly fitted survival models rather than independently fitted OS models based on the argument that the proportional hazards assumption is plausible after examining the log

cumulative hazard plots and undertaking a statistical test using the observed OS data. The company then used the HR for OS reported in the RESORCE study[6] to derive the OS model for the BSC group instead of the treatment effect estimated via the jointly fitted model. The ERG disagrees with the company's approach to model the OS data for several reasons.

Firstly, not all the parametric distributions fitted by the company belong to the family of parametric proportional hazards models. For example, the log normal, log logistic and gamma are AFT models and do not make assumptions of proportional hazards. It is not appropriate to apply an HR to an AFT model. Secondly, where applicable, the validity of the proportional hazards assumption in the observed period does not necessarily hold in the unobserved period; the clinical validity of the proportional hazards assumption should be assessed in the extrapolation period. Thirdly, the goodness-of-fit of the fitted OS curve in the BSC arm using the reported HR in the RESORCE study was not assessed. The AIC and BIC statistics shown in Table 20 and the observed and predicted OS curves shown in Figure 11 were generated using the treatment effect coefficient estimated from the jointly fitted models, not the reported HR. This inappropriate use of the reported HR may have an impact on the ICER for regorafenib versus BSC.

In response to a request for clarification on this matter (see company's clarification response,[8] question B2), the company presented the results of an analysis in which the OS curve for the BSC group was modelled using the log normal function including a constant acceleration factor derived from the jointly fitted model. In this scenario, the ICER is higher than the company's original base case (ICER=£37,239 per QALY gained).

In addition, the company's clarification response also included further analyses in which independent parametric curves were fitted to the available OS data; these independent models do not require a treatment effect covariate and do not impose restrictive assumptions about proportional hazards/odds between competing treatment groups. Within these analyses, the log normal curve resulted in the lowest AIC and BIC, although the AIC for the generalised gamma function was only slightly higher. The curve fits appear similar between the candidate OS functions.

Table 32 presents the results of the company's re-analysis of the OS data using the independently fitted models. As shown in the table, the ICER generated using log normal survival functions fitted independently to the OS data for each treatment group is similar to the ICER generated using the company's base case model.

Table 32: Company's original base case results and results generated using independently fitted OS curves (adapted from company's clarification response, question B2)

Scenario	Inc. costs	Inc. QALYs	ICER (per QALY gained)
Company's original base case - OS modelled using dependent log normal functions with HR treatment effect	£12,262	0.367	£33,437
Independently fitted OS models (no treatment effect covariate)			
Log logistic	£12,040	0.360	£33,463
Log normal	£12,276	0.368	£33,334
Weibull	£6,161	0.244	£25,248
Exponential	£7,670	0.290	£26,428
Generalised gamma	£12,438	0.377	£33,028
Gompertz	£7,165	0.265	£27,033

Inc. – incremental

(2) Limited consideration of clinical plausibility of extrapolated OS curves

The ERG notes that the discussion of clinical validity within the CS relates only to the differences between the observed and predicted OS estimates at cycle 35 (another measure of goodness-of-fit rather than plausibility), and differences between the extrapolated OS estimates derived from the exponential, log normal and log logistic models at the 5-year and 10-year timepoints. The CS does not contain any formal assessment of the clinical plausibility of the extrapolated survival times.

In contrast to the parametric model selected within the company's base case analysis, one clinical advisor to the ERG did not consider the log normal OS distribution to be clinically plausible. They noted that the model-predicted sustained gap in OS between the regorafenib and placebo groups beyond 35-cycles produced by the log normal function was unrealistic within the progressed HCC population. As a consequence, the advisor therefore considered the log logistic and generalised gamma functions to also be clinically implausible. The advisors' preferred curve was the Weibull function, although they noted that both the exponential and Gompertz functions were very similar and were therefore also potentially plausible. The ERG's second clinical advisor did not state a strong preference in favour of any of the individual parametric functions.

(3) Concerns regarding the modelling of time to treatment discontinuation to estimate regorafenib acquisition costs

The ERG has concerns regarding the approach taken to estimate the amount of regorafenib received over time. The company's model estimates time on treatment separately during the progression-free and post-progression phases. During the progression-free phase, the probability of receiving treatment is modelled according to the PFS curve and a compound probability of discontinuation (an additional [redacted] patients discontinue during each model cycle). The per-cycle probability of discontinuing regorafenib was estimated by dividing the proportion of patients who discontinued treatment for more

than one cycle prior to disease progression ([redacted]) by the median PFS duration in the regorafenib group (3.1 months). The probability of having discontinued regorafenib during each cycle whilst progression-free is calculated using the following equation:

Probability of having discontinued treatment at time t = *Probability of having discontinued treatment at time t-1 x (1+ per-cycle discontinuation probability)* [i]

The ERG does not believe that this approach is logically correct, but notes that setting this discontinuation rate equal to zero has only a minor impact on the cost-effectiveness of regorafenib (ICER = £33,749 per QALY gained).

During the post-progression phase, the company's model estimates the proportion of patients who have progressed and are still receiving regorafenib treatment. This is calculated using the post-progression treatment probability together with the sumproduct of the probability of being newly progressed in the given cycle and the post-progression treatment continuation rate. This approach assumes that the probability of receiving post-progression treatment and the post-progression treatment continuation rate are independent of the time at which the progression occurs. The ERG notes that this assumption may not be valid and the overall approach to modelling time on treatment is overly complex and makes unnecessary assumptions where data exist.

Given that [redacted] of patients continued to receive regorafenib treatment following disease progression, it is unclear why the company's model divides the total treatment received according to the presence or absence of disease progression. The ERG considers that the most appropriate approach to estimating the amount of drug received would instead involve the direct use of the time to treatment discontinuation (or death) curves observed within the RESORCE trial.[6] Such an approach would also render the company's approach to modelling pre-progression discontinuation redundant.

In response to a request for clarification, the company provided an analysis of time to treatment discontinuation within the regorafenib group of the RESORCE trial.[6] This analysis involved the consideration of two separate time to treatment discontinuation Kaplan-Meier curves. "Curve A" assumed that patients did not continue treatment beyond the 29th February 2016 DCO (see Figure 15). "Curve B" assumed that patients who were still receiving treatment on the 29th February 2016 were censored (see Figure 16). As indicated in the company's clarification response,[8] "Curve A" corresponds to the assumptions made in the company's original submitted model. The ERG notes that this approach assumes that all patients who are still on treatment at the DCO immediately discontinue regorafenib. This is clearly inappropriate, hence the subsequent analyses of "Curve A" are not considered further within this ERG report.

Figure 15: "Curve A" – Kaplan-Meier curve for time to treatment discontinuation assuming that patients discontinue treatment at the 29th February 2016 cut-off (reproduced from company's clarification response, question B8)

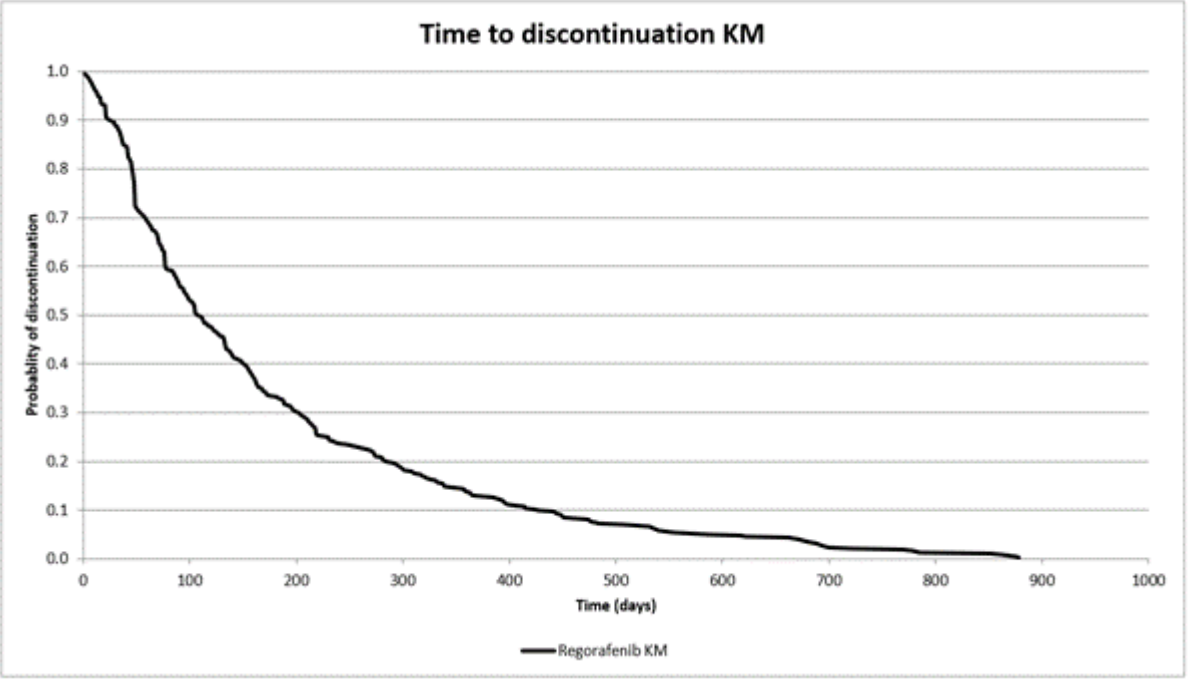
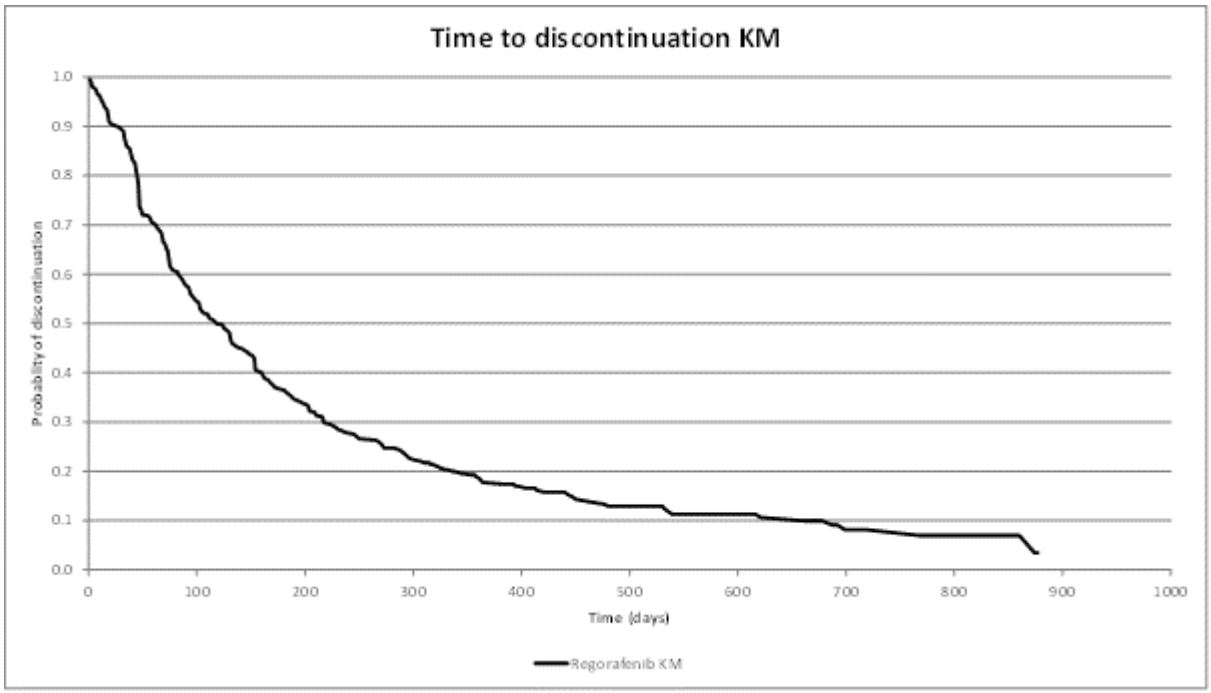


Figure 16: "Curve B" – Kaplan-Meier curve for time to treatment discontinuation assuming that patients on treatment on 29th February 2016 are censored (reproduced from company's clarification response, question B8)



Within their clarification response,[8] the company fitted parametric curves (log normal, log logistic, Weibull, exponential and Gompertz) to the available data on time to treatment discontinuation. The generalised gamma function was not considered; the company's clarification response does not explain this omission. Statistical goodness-of-fit of the candidate parametric models was considered through the use of AIC and BIC statistics (see Table 33). Plots of the fitted parametric curves are shown in Figure **17**. As shown in Table 34, the use of these parametric curves increases the ICER to the range £38,741 to £39,207 per QALY gained.

Table 33: AIC and BIC statistics for parametric models fitted to time to treatment discontinuation data, patients on treatment on 29th February 2016 censored (adapted from company's clarification response, question B8)

Distribution	AIC	BIC
Log logistic	**1145.59**	**1153.44**
Log normal	1148.23	1156.07
Gompertz	1159.86	1167.70
Weibull	1176.79	1184.63
Exponential	1179.11	1183.03

Lowest values highlighted in bold

Figure 17: Parametric models – time to treatment discontinuation, patients on treatment on 29th February 2016 censored (reproduced from company's clarification response, question B8)

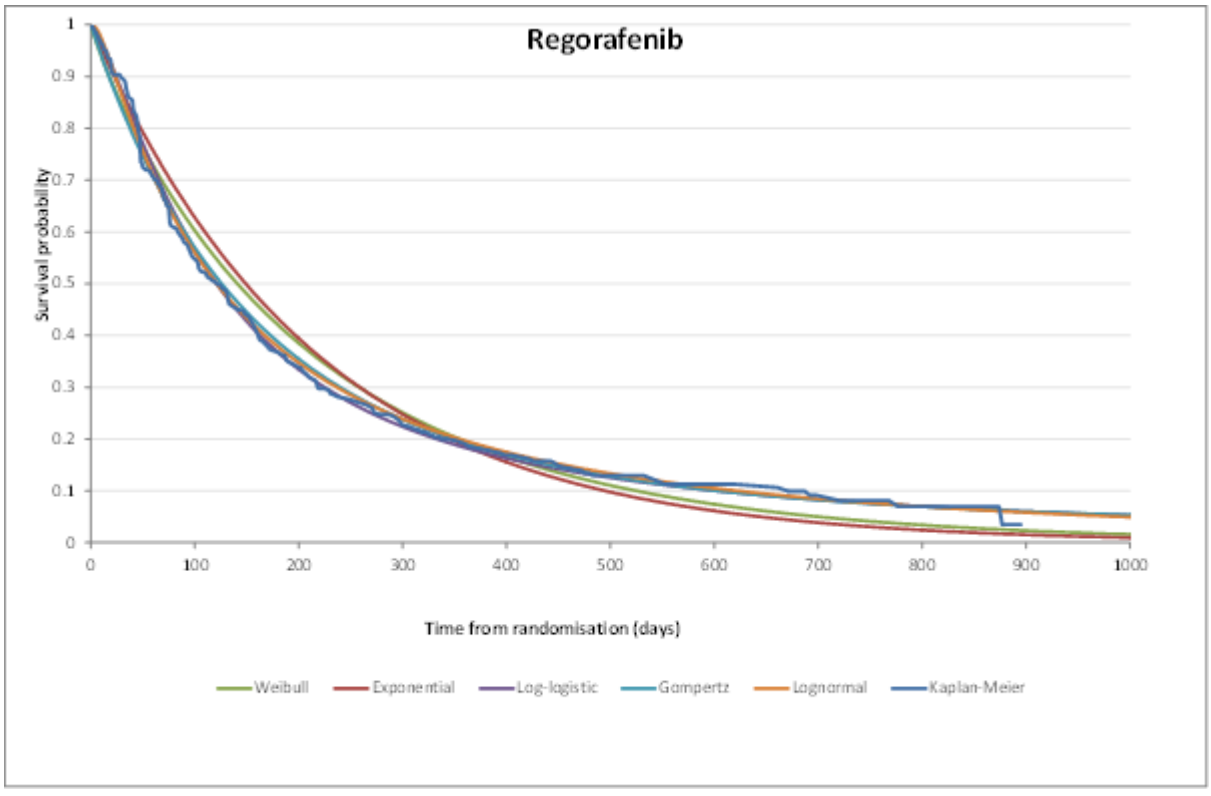


Table 34: Cost-effectiveness results for alternative curves fitted to time to treatment discontinuation, patients on treatment on 29th February 2016 censored (adapted from company's clarification response, question B8)

Time to treatment discontinuation scenario	Incremental costs (regorafenib versus BSC)	ICER (per QALY gained)
Original base case	[redacted]	£33,437
Raw KM treatment data	[redacted]	£38,906
Log normal	[redacted]	£39,207
Log logistic	[redacted]	£38,741
Weibull	[redacted]	£38,985
Exponential	[redacted]	£38,905
Gompertz	[redacted]	£39,060

(4) Inclusion of potentially unrealistic cost savings due to dose reductions/interruptions

The company's model includes cost savings associated with dose reductions and treatment interruptions for regorafenib. One clinical advisor to the ERG stated that should regorafenib be made available on the NHS, it would be prescribed monthly according to a fixed delivery schedule. The clinical advisor also noted that the logistics of current prescribing practices in their centre do not allow for the reduced frequency of individual prescriptions for patients with leftover pills; rather, any pills not taken by the patient would be returned and destroyed. Consequently, the ERG does not believe that the cost reductions included in the company's model would be fully realised in clinical practice and instead has costed regorafenib costed at its full maximum dose of 160mg per day for the entire duration of treatment within the exploratory analyses (see Section 5.5). As shown in the company's DSAs, the inclusion of full treatment costs increases the ICER for regorafenib versus BSC considerably (ICER excluding dose reductions = £41,206 per QALY gained, see Table 29). The ERG acknowledges that where the reduction in dose is planned and a lower dose is to be maintained in the long-term, the ERG's assumption of 160mg per day for each patient will overestimate the ICER for regorafenib.

(5) Concerns regarding expert clinician survey to inform health state resource use

Within the CS,[1] the company refers to a survey conducted in 2015 with three "*leading clinical experts in the field of oncology in the UK*" that was undertaken to estimate resource use associated with sorafenib and for patients receiving BSC. The company assumed that the sorafenib results were generalisable to regorafenib, although the CS notes that there is currently no experience in the clinical setting with regorafenib in the treatment of HCC. The CS does not make reference to an earlier survey which was conducted in 2007 using four UK clinicians, despite the fact that within the earlier sorafenib appraisal,[5] the NICE Cancer Drugs Fund (CDF) Appraisal Committee preferred the pooled analysis of both the 2007 and 2015 surveys.

For both the CDF appraisal of sorafenib and the clarification questions relating to regorafenib, the company have stated that the 2017 survey is preferable as "*The estimates from 2007 precede the availability of sorafenib and are not based on clinical experience. In contrast the estimates from 2015 are based on clinician experience in the use of sorafenib since its launch in 2008*" (company's clarification response,[8] question B14).

In the sorafenib CDF appraisal, the NICE Decision Support Unit (DSU) expressed a contrary view, stating that: "*The DSU thinks that discarding the results of the original survey is not the best option, especially considering that the original survey involved more clinicians and contained more responses… The estimates of the clinicians that took part in the new survey might have produced better estimates for the sorafenib arm due to the learning curve but the estimates for the BSC arm from the original survey should be equally valid when compared with those of the new survey.*"[43]

Table 35 summarises the completion rates for the 2007 and 2015 surveys.

Table 35: Comparison of the number of responses collected in the 2007 survey compared with the 2015 survey (adapted from DSU report on sorafenib[43])

	2007 survey	**2015 survey**
Total number of questions	279	247
Questions with no responses (%)	39 (14.0)	16 (6.5)
Questions with one responses (%)	31 (11.1)	35 (14.2)
Questions with two responses (%)	33 (11.8)	100 (40.5)
Questions with three responses (%)	36 (12.9)	96 (38.9)
Questions with four responses (%)	140 (50.2)	0 (0.0)
Total responses	765	523
Average number of responses	2.74	2.12

In the factual accuracy check round for the sorafenib CDF appraisal, the company stated that a preference for the 2015 survey "*on the grounds that health technologies and resource use change over time*" should be made. The DSU (acting as an ERG) responded stating: "*The ERG notes that the difference between the estimates of the physicians taking part in the survey points to uncertainty rather than changes in best supportive care (BSC). For example, in the new survey, the percentage* [redacted] *was estimated to be* [redacted] *by the first physician and* [redacted] *by the second (the third physician's estimate is not available). Similarly, the number* [redacted] *by the first physician was* [redacted] *and* [redacted] *by the second physician (the third physician's estimate is not available). These two parameters are the two main drivers of the difference between the ICERs using the old and new resource use estimates. The ERG believed including the estimates of the 4 physicians that took part in the original survey resulted in more robust estimates."*[43]

The ERG for this appraisal (of regorafenib) notes that there are no new data presented which would alter the judgment of the CDF Appraisal Committee. As such, and noting the arguments put forward by the DSU, the ERG maintains that the pooled estimates are preferable to the 2015 survey responses alone.

For the sake of clarity, the ERG has tabulated the resource use estimates taken from the 2015 survey and the pooled 2007 and 2015 surveys (see Table 36). It should be noted that monthly estimates have been assumed to be generalisable to 28-day cycles. These data are conditional on whether a patient is on treatment and whether the patient is in a pre-progression or post-progression state. It is observed that regardless of which survey responses are used, the rates of patients requiring hospitalisation were lower for those on regorafenib as were the assumed durations of hospital stays and thus the cost per hospitalisation is lower. Clinical advice received by the ERG indicates that it is plausible that the use of regorafenib could reduce the number of hospitalisations compared with BSC alone.

It should also be noted that a potential discrepancy was found in the survey data and the way in which these were interpreted and implemented by the company. Further details are provided in Appendix 1 however, briefly, it appears that patients requiring hospitalisation are assumed to have fewer than one hospital visit per month on average, which is not logical. The company states that they had assumed *a priori* that this number would be one or greater, which the ERG believes is logical. Appendix 1 contains a replication of the company's response, which attempts to justify the data used in the CS, and a sensitivity analysis performed by the company in which the number of hospitalisations per month for those requiring hospitalisation is set to one. The ERG does not accept the justification put forward by the company and prefers the assumptions used in the sensitivity analyses performed by the company.

The ERG considers that there are still implementation errors in non-hospital costs within the CS when data from the pooled survey are used, but that the correction of these will have only a minor impact on the ICER and thus have left these at the values used by the company.

Table 36: Assumed resource use and costs per 28-day treatment cycle

Resource item	Unit cost	Progression-free - proportion using resource				Post progression - proportion using resource			
		Pooled 2007 & 2015 surveys		2015 survey only		Pooled 2007 & 2015 surveys		2015 survey only	
		Sorafenib*	BSC	Sorafenib*	BSC	Sorafenib*	BSC	Sorafenib*	BSC
Hospitalisation									
General ward	£801								
Duration of stay (days)	-								
A&E admission	£138								
Estimated cost per hospitalisation [1]	-								
Hospital outpatient appointments									
Oncologist	£163								
Hepatologist	£253								
Gastroenterologist	£132								
Clinical nurse specialist	£130								
Palliative care team	£131								
Radiologist	£135								
Macmillan nurse	£73								
Follow up visits									
GP visit	£36								
Nurse visit	£36								
Specialist visit	£151								
Tests									
AFP	£3.03								
Liver function	£2.78								
Biochemistry	£1.34								
Complete blood count	£2.65								
International normalised ratio	£3.43								
Endoscopy	£743								
Radiological tests									
CT scan of abdomen	£122								
MRI of abdomen	£238								

** Assumed to apply to regorafenib; [1] Calculated multiplying the estimated length of stay by the estimated cost of a bed day on a general ward (£801); [2] 0.93 at progression; [3] 0.78 at progression; [4] 0.60 at progression; [5] 0.04 at progression; [6] 1.00 at progression; [7] 0.67 at progression*

(5) Likely overestimation of the cost of a general ward bed day

The company's model includes the cost of a general ward bed day of £801: this estimate was derived from a response to a Freedom of Information Act request. According to the CS,[1] this reflects the fully absorbed cost. No further details of the derivation or source of this value are presented within the CS and it is unclear why current NHS Reference Costs[37] have not been used (as was done within the previous sorafenib appraisal). The ERG notes that based on the NHS Reference Costs 2015/16,[37] for non-elective long-stay admissions, the mean cost per bed day weighted by the number of total finished consultant episodes (FCEs) is £572.44. This estimate is lower than the unit cost applied within the company's model.

(6) Use of potentially inappropriate NHS Reference Costs

The ERG notes that some of the costs included in the company's model may not reflect the best use of the available NHS Reference Costs. These are detailed below.

1. The company's model assumes that the cost of an A&E admission is £138.00, based on the total number of FCEs. The ERG notes that the weighted mean cost for patients admitted to A&E excluding episodes relating to emergency dental work and patients who are dead on arrival is £204.11 per episode. The ERG believes that this represents a more appropriate unit cost.
2. The company's model uses a cost of £131.00 for a palliative care team visit; the CS states that this is based on the follow-up cost for a face-to-face consultant-led follow up outpatient appointment for pain management contained in the NHS Reference Costs 2015/16.[37] However, the ERG considers that it may be more appropriate to use a weighted average of outpatient palliative pain management costs (healthcare resource group [HRG] codes SD04A & SD05A); this corresponds to a weighted average of £119.03 per visit. The ERG acknowledges that it asked the company to use a cost of £131 in the clarification process.
3. The company uses a cost £151.12 per specialist follow-up visit. The ERG was unable to identify this value within the NHS Reference Costs 2015/16.[37] The ERG considers that the tariff cost for a medical oncology consultant-led, non-admitted face-to-face visit would be more appropriate (cost=£162.84).
4. The model assumes a cost of £238.00 for an abdominal MRI scan based on HRG code RD03Z. The ERG was unable to find this value within the NHS Reference Costs 2015/16.[37] The ERG believes that the most appropriate tariff value is that of an outpatient MRI scan (cost=£202.70).
5. The company's costing of AEs uses an unweighted average for the costs applicable to each type of AE. The use of weighted mean costs changes the overall AE cost (for those experiencing AEs) to £1,184.11 for regorafenib and £1,365.07 for BSC alone.

(7) Questionable reliability of post-progression utility estimate

The ERG has two concerns regarding the health utilities included in the company's model. These relate to: (i) the questionable reliability of the post-progression utility estimate, and (b) the potential underestimation of the impact on regorafenib treatment on HRQoL.

(a) Questionable reliability of post-progression utility estimate

The ERG has doubts about the face validity of the utility values collected in the RESORCE trial[6] as the utility decrement associated with progression was only -0.048. This point was raised with the company at the clarification stage[8] (question B11). In response, the company stated that: "*We consider the values derived from the RESORCE study to be face-valid.*" The company also stated that the cost-effectiveness of regorafenib was found to be relatively insensitive to doubling the decrement associated with progression (see CS,[1] Table 58).

The ERG believes that the central estimate of the disutility associated with progression is likely to represent an underestimate. As shown in Table 21, the EQ-5D response rate for patients in the pre-progression state was high (typically greater than 90%) and is thus representative of patients in the RESORCE trial,[6] although the ERG notes that the estimated pre-progression EQ-5D score of 0.811 appears high for a population with advanced HCC who have previously progressed on sorafenib. The percentage of patients in the post-progression state completing the EQ-5D was much lower, typically between 20% and 30%, which raises the possibility that only the patients in the best health at that time point completed the EQ-5D questionnaire.

(b) Potential underestimation of utility decrements associated with regorafenib treatment

Within the RESORCE trial,[6] the EQ-5D questionnaire was provided on the first day of each treatment cycle, when a patient had previously spent a week without treatment. The EQ-5D assesses the respondent's health at the time of completion and does not consider patients' health over previous days or weeks. As such, any deleterious effects of regorafenib treatment may not be captured due to the timing of administration of the EQ-5D.

(8) Insufficient consideration of uncertainty

The ERG notes that the company's PSA includes a number of parameters which are held fixed. These include: (a) the PFS curves; (b) the baseline OS curve for the regorafenib group; (c) the post-progression treatment continuation rates, and (d) the unit costs associated with health state resource use (see Table 27). These are all uncertain variables and as such they should have been included in the company's PSA. Consequently, the company's PSA underestimates the uncertainty surrounding the incremental costs and effects of regorafenib. In response to a request for clarification[8] (question B5), the company amended the model to include uncertainty surrounding the PFS curves. However, the other time-to-

event curves (OS and time to treatment discontinuation) remain fixed within the PSA. This may be important as the use of log normal distributions (as used to model OS) tend to increase probabilistic ICERs relative to their deterministic counterparts. Owing to time constraints, the ERG did not modify the company's uncertainty analysis.

5.4 Model amendments and revised base case submitted following clarification

Following the clarification process, two further amended versions of the model were submitted by the company.

Post-clarification model 1: Company's revised model with increased functionality and corrections using RESORCE 29th February 2016 data cut-off

Following the clarification process, the company submitted a revised model that addresses some of the issues identified within this critical appraisal. The key features of this revised model relate to additional functionality to select:

(i) Separate parametric OS curves fitted independently to each treatment group (without a treatment effect covariate)

(ii) The treatment effect covariate generated from the jointly fitted model rather than the Cox-derived HR

(iii) The use of a palliative care team visit cost of £131 (rather than £136, as applied in the original model)

(iv) The use of 13.044 cycles per year (rather than 13.00, as applied in the company's original model)

(v) The correction of the programming error in which the QALY loss for AEs in the BSC group was erroneously linked to the regorafenib AE rate

(vi) The resource use estimates derived from the earlier clinical survey used in the previous sorafenib appraisal (2007 and 2015 surveys combined)[5]

(vii) The use of parametric time to treatment discontinuation curves to estimate drug acquisition costs. A further option is enabled which allows this analysis to be based on all patients on treatment discontinuing at the 29th February DCO or being censored at this DCO. The ERG notes however that the company incorrectly truncated the total treatment costs at 29 cycles, thereby ignoring additional costs incurred due to the tail of the curve.

A single preferred base case analysis was not presented using this version of the model.

Post-clarification model 2: Company's revised base case model using RESORCE 23rd January 2017 data cut-off

Subsequent to the submission of the first post-clarification model, the company submitted a further revised model and revised base case analysis which incorporated updated time-to-event data. The company investigated the possibility of using a longer follow-up period with adjustments made for patients who had started on BSC who subsequently received regorafenib treatment. However, as only four of the 194 patients (2.1%) initially randomised to receive BSC received regorafenib the company did not perform statistical adjustment for treatment switching '*as this represents such a small proportion of patients*'. The ERG notes that this will be unfavourable to regorafenib. However, the company only analysed dependent OS curves including a treatment effect covariate. The company's clarification response[8] (question B1) provides results using this version of the model which include some of the adjustments above (including the use of parametric time to treatment discontinuation curves), but does not, however, have the functionality to apply independently fitted OS curves for each treatment group '*due to time constraints'*; this makes these results difficult to interpret. The key features of the company's revised base case analysis are:

- Use of 23rd January 2017 DCO
- PFS modelled using observed Kaplan-Meier curves
- OS modelled using dependent log normal functions using a revised treatment effect covariate (note: it is unclear whether this is a revised HR or the jointly fitted model treatment effect covariate)
- Time to treatment discontinuation modelled using log logistic function (note: the error relating to the truncation of this curve at 29 cycles also applies within this model version).
- Utilities based on RESORCE trial[6] (as per the original CS[1])
- Correction of the costs of palliative care team visits, the number of cycles per year and the BSC AE rate programming error (see clarification response,[8] questions B16, B22 and B24)
- Use of the 2015 sorafenib resource use survey (as per the original CS[1])

The company's revised deterministic base case results are presented in Table 37: these results also include the correction of a further programming error identified by the company. The company's revised base case analysis leads to a slightly higher ICER compared with their original model: the deterministic ICER for regorafenib versus BSC is estimated to be £36,050 per QALY gained.

Table 37: Company's revised base case cost-effectiveness results – regorafenib versus BSC

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Regorafenib	1.073	[redacted]	0.406	£14,625	**£36,050**
BSC	0.668	[redacted]	-	-	-

Inc. – incremental

The company's clarification response[8] (question B1) and the further revised analyses received following the company's identification of the programming error include additional scenario analyses using their revised base case model. However, as the ERG has concerns which are not reflected in the company's revised base case model (in particular, the inappropriate use of dependent OS curves and the use of an erroneously truncated time to discontinuation curve), the results of these additional scenario analyses are not presented here.

5.5 ERG's exploratory analyses

5.5.1 ERG's exploratory analyses - methods

The ERG undertook seven sets of exploratory analyses. All analyses were undertaken using the deterministic version of the revised model submitted by the company following clarification, which uses the 29th February 2016 DCO (post-clarification model 1). It was not possible to incorporate all of the ERG's preferred assumptions using the model which incorporates data from the January 23rd DCO (post-clarification model 2), hence all exploratory analyses are limited in this respect. Additional sensitivity analyses were undertaken using the ERG's preferred base case scenario. These include the exploration of alternative parametric functions for OS and time to treatment discontinuation, alternative assumptions regarding HRQoL, an alternative interpretation of the resource use survey data, and the optimistic assumption of cost savings associated with a sustained mean daily dose of 120mg regorafenib.

Exploratory analysis 1: Correction of unequivocal model errors and use of alternative unit costs

The following corrections were made to the company's revised model:

- The number of cycles per year was set equal to 13.044*
- The programming error relating to the AE rate for the BSC group was corrected*
- The programming errors relating to the hospitalisation and medical visit costs for patients during the progression-free phase were corrected
- The proportion of BSC patients receiving post-progression regorafenib was set equal to zero.

** These changes were added by the company as options after the clarification process*

In addition, the following unit costs were amended as follows:

- The cost per A&E visit was set equal to £204.11.
- The cost per palliative care team visit was set equal to £119.03.
- The cost per specialist follow-up visit was set equal to £162.84
- The cost per abdominal MRI scan was set equal to £202.70.
- The cost of each AE were set equal to £1,184.11 for the regorafenib group and £1,365.07 for the BSC group.

All subsequent exploratory analyses include these corrections and amendments.

Exploratory analysis 2: Inclusion of more appropriate general ward bed day cost

The general ward bed day cost was amended to reflect the weighted cost of an excess bed day (£572.44 per bed day).

Exploratory analysis 3: Use of full pack dosing

Cost savings due to reduced dosing and treatment interruptions were removed from the model: this was implemented by setting the mean daily dose of regorafenib equal to 160mg per day.

Exploratory analysis 4: Removal of half-cycle correction for drug acquisition costs

Total drug costs were calculated using the proportion of patients alive and on treatment at the beginning of each model cycle. The use of half-cycle correction was retained for all other cost and QALY calculations.

Exploratory analysis 5: Use of combined 2007 and 2015 survey costs

The results of the pooled 2007 and 2015 surveys were used to inform health state resource use. To address the apparent logical inconsistencies in the results of the surveys, it was assumed that the proportions of patients requiring hospitalisation for those on regorafenib and BSC were correct and that these patients were only hospitalised once per month. This is the same approach used by the company in their sensitivity analyses provided post-clarification.

Exploratory analysis 6: Use of independent Weibull functions to model OS

The model was amended to use the independent Weibull functions for OS (excluding a treatment effect covariate).

Exploratory analysis 7: Use of a fully extrapolated log logistic time to discontinuation curve (patients on treatment at 29th February 2016 censored)

In line with the company's revised base case, the log logistic model was selected to model time to treatment discontinuation, based on the time-to-event data which includes censoring of patients remaining on treatment at the 29th February 2016 DCO. A new worksheet was added by the ERG which estimates discounted drug acquisition costs per cycle based on the log logistic function. The worksheet includes full extrapolation up to 10-years (the last timepoint from the company's parametric curve-fitting). A logical consistency constraint was also added to ensure that the probability of being alive and on treatment could not be greater than the survival probability predicted by the selected OS curve. This analysis also includes the assumption of full pack dosing.

Exploratory analysis 8: ERG's preferred base case

The ERG's preferred base case includes exploratory analyses 1 to 7.

5.5.2 Results of the ERG's exploratory analyses

Table 38 presents the results of the ERG's exploratory analyses.

Table 38: Results of the ERG's exploratory analyses

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Company's base case (revised base case model, deterministic)					
Regorafenib	1.073	[redacted]	0.406	£14,625	**£36,050**
BSC	0.668	[redacted]	-	-	-
Exploratory analysis 1: Correction of unequivocal model errors and use of alternative unit costs					
Regorafenib	1.048	[redacted]	0.368	£12,659	**£34,406**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 2: Inclusion of more appropriate general ward bed day cost**					
Regorafenib	1.048	[redacted]	0.368	£12,647	**£34,376**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 3: Use of full pack dosing**					
Regorafenib	1.048	[redacted]	0.368	£15,508	**£42,151**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 4: Removal of half-cycle correction for drug acquisition costs**					
Regorafenib	1.048	[redacted]	0.368	£13,332	**£36,235**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 5: Use of combined 2007 and 2015 survey costs**					
Regorafenib	1.048	[redacted]	0.368	£20,297	**£55,166**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 6: Use of independent Weibull functions to model OS**					
Regorafenib	0.896	[redacted]	0.265	£10,242	**£38,683**
BSC	0.632	[redacted]	-	-	-
*Exploratory analysis 7: Use of a fully extrapolated log logistic time to discontinuation curve (patients on treatment at 29th February 2016 censored, with full pack dosing)**					
Regorafenib	1.048	[redacted]	0.368	£21,751	**£59,120**
BSC	0.680	[redacted]	-	-	-
*Exploratory analysis 8: ERG's preferred base case (including all individual amendments)**					
Regorafenib	0.896	[redacted]	0.265	£21,468	**£81,081**
BSC	0.632	[redacted]	-	-	-

Inc. – incremental

** Exploratory analyses 2-8 also include corrections and amendments made in exploratory analysis 1*

The correction of model errors and the use of more appropriate unit costs does not have a marked impact upon the ICER for regorafenib versus BSC (ICER=£34,406 per QALY gained). In addition, the use of a lower cost per hospital bed day also has only a marginal impact upon the ICER (ICER=£34,276 per QALY gained). The removal of the half-cycle correction of acquisition costs, the full costing of all prescribed packs of regorafenib, the use of the pooled 2007 and 2015 resource use surveys and the use of a fully extrapolated log logistic function to estimate time to treatment discontinuation each

individually result in a less favourable ICER relative to the company's base case scenario (ICER range = £36,235 to £59,120 per QALY gained). The use of independent Weibull functions to model OS also increase the ICER for regorafenib versus BSC (ICER=£38,683 per QALY gained). The ERG's preferred base case, which includes all of the above amendments (exploratory analyses 1 to 7), results in an ICER for regorafenib versus BSC of £81,081 per QALY gained. The ERG notes that this base case includes data from the 29th February 2016 DCO of the RESORCE study: this is because the company's revised base case model which uses the January 23rd 2017 DCO does not include the functionality for modelling independently fitted OS curves. The ERG prefers the use of independent curves and believes that the company's revised base case which uses the later DCO were less appropriate due to the use of dependent curves. The ERG's base case ICER for regorafenib using both the ERG's preferred assumptions and the later DCO of the RESORCE study is unknown and cannot be assessed using the available versions of the company's model.

5.5.2 Additional sensitivity analyses using the ERG's preferred base case model

Table 39 presents additional sensitivity analyses undertaken using the ERG's preferred base case.

Table 39: Additional sensitivity analyses undertaken using ERG preferred base case model

Scenario	Inc. QALYs	Inc. costs	ICER (regorafenib versus BSC)
ERG base case	0.265	£21,468	£81,081
Alternative OS functions			
OS - exponential	0.311	£22,690	£72,959
OS – log normal	0.369	£27,617	£74,744
OS – log logistic	0.361	£27,363	£75,792
OS – Gompertz	0.286	£20,757	£72,642
OS – generalised gamma	0.378	£27,893	£73,826
Alternative time to treatment discontinuation functions			
TTTD - exponential	0.265	£19,625	£74,122
TTTD – Weibull	0.265	£19,942	£75,317
TTTD – log normal	0.265	£21,606	£81,602
TTTD – Gompertz	0.265	£21,633	£81,703
Alternative utility values			
Utilities from SHARP trial	0.232	£21,468	£92,719
Disutility due to progression doubled (state utility=0.715)	0.260	£21,468	£82,689
Disutility due to progression tripled (state utility=0.667)	0.254	£21,468	£84,362
Alternative interpretation of company's resource use survey			
Number of hospitalisations per month estimated per month assumed to apply to the entire population.	0.265	£22,006	£83,114
Inclusion of dose reductions			
Indefinite dose reduction to 120mg/day	[redacted]	[redacted]	[redacted]

Inc. – incremental

The results of the ERG's additional sensitivity analyses indicate that alternative choices of parametric functions to model OS may reduce the ICER for regorafenib (ICER range = £72,642 to £81,081 per QALY gained). The use of alternative parametric functions to model time to treatment discontinuation leads to ICERs in the range £74,122 to £81,703 per QALY gained. The use of the utilities from the SHARP trial increase the ICER for regorafenib versus BSC to £92,719 per QALY gained. Increasing the disutility associated with progressed disease (relative to the progression-free utility score) does not have a substantial impact on the ICER for regorafenib. The exploratory analysis in which the number of hospitalisations per month estimated in the survey was applied to the entire population has only a minor impact on the ICER for regorafenib compared with assuming that the percentage requiring hospitalisation was correct and that patients were hospitalised once per month. The inclusion of dose reductions to [redacted] for all patients from the start of treatment reduces the ICER to [redacted] per QALY gained; the ERG notes that this represents a highly optimistic scenario and that the ICER for regorafenib is likely to be higher than this estimate.

5.6 Discussion

The CS includes a systematic review of published economic evaluations of treatments for HCC together with a *de novo* health economic evaluation of regorafenib (plus BSC) versus BSC alone in patients with HCC. The company's review did not identify any economic evaluations of regorafenib within this indication. Additional searches undertaken by the ERG identified one economic evaluation study which assessed regorafenib versus BSC in patients (Parikh *et al*[35]); this study was published after the company's searches had been carried out. The company and the ERG both agreed that this study is not relevant to the current appraisal due to the use of a short time horizon, the absence of any form of extrapolation of time-to-event outcomes and the use of a US health care system perspective.

Owing to the absence of any relevant existing studies, the company developed a *de novo* partitioned survival model to assess the cost-effectiveness of regorafenib (plus BSC) versus BSC alone in adult patients with unresectable HCC who have been previously treated with sorafenib. Incremental health gains, costs and cost-effectiveness of regorafenib are evaluated over a 15-year time horizon from the perspective of the NHS and PSS. The company's model includes three health states: (1) progression-free; (2) progressed disease, and (3) dead. The model parameters were mostly informed by analyses of time-to-event data (PFS, OS and time on treatment) collected within the RESORCE trial[6] (January 29th 2016 DCO). PFS was modelled using the observed PFS estimates, OS was modelled using a log normal distribution with a treatment effect covariate (an HR) and time to treatment discontinuation was modelled using a "cycle-cohort simulation" approach. Resource use was informed by a survey of three clinical experts undertaken in 2015. The model assumes that a small proportion of patients treated with regorafenib will discontinue prior to disease progression and that a proportion of patients continue

regorafenib treatment following progression. The model includes a mean daily dose of regorafenib which accounts for dose reductions and treatment interruptions observed within the RESORCE trial.

Based on the probabilistic version of the company's original model (assuming the log normal function for OS), regorafenib is expected to generate an additional 0.37 QALYs at an additional cost of £12,311: the corresponding ICER for regorafenib versus BSC is £33,335 per QALY gained. The deterministic version of the company's base case model produces a very similar ICER of £33,437 per QALY gained. Assuming a willingness-to-pay (WTP) threshold (λ) of £30,000 per QALY gained, the company's model indicates that the probability that regorafenib produces more net benefit than BSC is 0.21. Assuming a WTP threshold of £50,000 per QALY gained, the probability that regorafenib produces more net benefit than BSC is 1.0.

The ERG critically appraised the company's economic analysis and double-programmed the deterministic version of the company's model. The ERG's critical appraisal identified a number of issues relating to the company's economic analysis and the evidence used to inform it. The most pertinent of these include: (i) the inappropriate use of an HR to model relative treatment effects on OS; (ii) limited consideration of the clinical plausibility of the extrapolated OS curves; (iii) concerns regarding the modelling of time to discontinuation of regorafenib; (iii) the inclusion of potentially unrealistic cost savings due to dose reductions and treatment interruptions; (iv) the use of the 2015 survey of three experts to inform health state resource use (and the exclusion of the earlier survey used to inform the earlier sorafenib appraisal); (v) concerns regarding the appropriateness of several unit cost estimates; (vi) the questionable reliability of the post-progression utility estimate and (vii) the inadequate representation of parameter uncertainty.

Following the clarification process, the company submitted two further versions of the model: (i) a revised model which includes additional functionality to address some of the issues identified within the ERG's critical appraisal, and (ii) a revised base case model which includes less functionality but uses the latest January 23rd 2017 DCO of the RESORCE trial. The company's revised base case analysis leads to a slightly higher deterministic ICER for regorafenib versus BSC of £36,050 per QALY gained compared with their original submitted model. Given that additional issues were identified by the ERG after receipt of this revised model (for example, the use of dependent OS curves and an erroneously truncated time to discontinuation curve), the ERG suggests that the results produced from this iteration of the company's model are not useful for informing decision-making.

The ERG undertook seven sets of exploratory analyses using the deterministic version of the company's revised model (using the 29th February 2016 DCO of the RESORCE trial). The ERG's preferred base case includes the following amendments: (i) the correction of model errors and use of alternative unit

costs; (ii) the inclusion of a more appropriate general ward bed day cost; (iii) the use of full pack dosing (no cost savings due to dose reductions or treatment interruptions); (iv) the removal of half-cycle correction for regorafenib acquisition costs; (v) the use of the combined 2007 and 2015 survey resource use estimates; (vi) the use of independent Weibull functions to model OS, and (vii) the use of a fully extrapolated log logistic time to discontinuation curve (patients on treatment at 29th February 2016 censored). The ERG's preferred base case, which includes all of these amendments, results in a deterministic ICER for regorafenib of £81,081 per QALY gained compared with BSC. The ERG notes that this ICER will be higher if a greater disutility associated with progression is assumed within the model. It should also be noted that where a reduction in dose is planned and the lower dose is to be maintained over the long-term, the ERG's assumption of indefinite full pack dosing for all patients will lead to an overestimation of the ICER for regorafenib. The additional sensitivity analyses undertaken by the ERG indicate that even under the highly optimistic assumption that all patients have indefinite dose reductions to [redacted] from the start of treatment, the ICER for regorafenib versus BSC remains above [redacted] per QALY gained. It is probable that the company will have information relating to whether dose reductions were due to clinically-planned reductions or due to other reasons: having this information would allow a more accurate estimation of the ICER.

6 END OF LIFE

NICE end of life supplementary advice should be applied in the following circumstances and when both the criteria referred to below are satisfied:

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment.

In the company's base case (Table 31) the mean expected life years associated with the use of BSC was estimated to be 0.90 years (10.8 months). This is markedly lower than stated 24-month cut-off. In the company's base case, regorafenib treatment was associated with a mean extension of life of 0.52 years (6.24 months) which is in excess of the stated 3-month cut-off. The changes made by the ERG relating to the choice of parametric OS functions do not change the conclusion with respect to the end of life criteria.

7 OVERALL CONCLUSIONS

The company's systematic review was generally well conducted. The review included a single, high-quality RCT: the RESORCE trial, which represents the relevant evidence. The trial reported that regorafenib was significantly more effective than placebo across the primary (OS) and secondary (PFS, TTP, ORR) outcomes, but also found that HRQoL was consistently worse on treatment than on placebo across different measures. AEs were frequent. The principal issue with the evidence concerns the limits of the trial population and how far they reflect the population seen in clinical practice in England. The RESORCE trial did not include some groups of adult HCC patients covered by the NICE scope and the licence, that is, those who are intolerant to sorafenib, or who are Child-Pugh class B or ECOG PS 2. The efficacy and safety of regorafenib in these groups is therefore uncertain.

The exploratory analyses undertaken by the ERG increase the ICER for regorafenib versus BSC from £36,050 per QALY gained (the company's revised base case) to an ERG-preferred ICER of £81,081 per QALY gained. The ERG notes that this ICER would increase slightly if a higher disutility for progressed disease is assumed. Where a reduction in dose is planned and the lower dose is to be maintained over the long-term, the ERG's assumption of indefinite full pack dosing for all patients will lead to an overestimation of the ICER for regorafenib. However, additional sensitivity analyses undertaken by the ERG indicate that even under the highly optimistic assumption that all patients have indefinite dose reductions to [redacted] from the start of treatment, the ICER for regorafenib versus BSC remains above [redacted] per QALY gained. Key differences in assumptions between the ERG and the company relate to: (1) the use of a fully extrapolated log logistic function to model time to treatment discontinuation; (2) the anticipated number of hospitalisations per month for those receiving regorafenib and for those on BSC, and (3) whether the acquisition costs of regorafenib pills not taken by a patient could be recouped.

7.1 Implications for research

The resource use, in particular frequency of hospitalisation, for patients on BSC after sorafenib should be recorded. If preferential rates are to be assumed for regorafenib, this should come from a large robust survey. If possible, the utility associated with patients who have progressed following treatment with regorafenib and with BSC should be estimated more robustly than was done in the follow-up of patients in the RESORCE study. Long-term data on OS and time to treatment discontinuation would improve the accuracy of the estimate of the cost-effectiveness of regorafenib.

8 REFERENCES

1. Bayer PLC Ltd. Regorafenib for previously treated unresectable hepatocellular carcinoma [ID991]. Company submission to the National Institute for Health and Clinical Excellence. Bayer: Reading, UK; 2017:1-179.
2. Cancer Research UK. Liver Cancer Statistics in the UK 2017. 2017. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer#heading-Zero (Accessed 18/01/2017).
3. European Association For The Study Of The Liver, European Organisation For Research And Treatment Of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *Journal of Hepatology* 2012;56:908-43.
4. Cholongitas E, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK. Systematic review: the model for end-stage liver disease – should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? *Alimentary Pharmacology & Therapeutics* 2005;22:1079-89.
5. National Institute for Health and Care Excellence. Sorafenib for treating advanced hepatocellular carcinoma (Cancer Drugs Fund reconsideration of TA189) - Final Appraisal Determination. London; 2017.
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. *The Lancet* 2017;389:56-66.
7. PRISMA. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA); 2009 (available from: http://www.prisma-statement.org/).
8. Bayer PLC Ltd. Regorafenib for previously treated unresectable hepatocellular carcinoma [ID991]. Response to ERG clarification questions. Bayer: Reading, UK;2017.
9. Bayer PLC Ltd. Regorafenib for previously treated unresectable hepatocellular carcinoma [ID991]. Company submission to the National Institute for Health and Clinical Excellence - Appendices. Bayer: Reading, UK; 2017:1-364.
10. Higgins J, Altman D, Gøtzsche P, Jüni P, Moher D, Oxman A, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
11. Lundh A, Lexchin J, Mintzes B, Schroll J, Bero L. Industry sponsorship and research outcome. *Cochrane Database of Systematic Reviews* 2017:Art. No.: MR000033. DOI: 10.1002/14651858.MR000033.pub3.
12. Stamatakis E, Weiler R, JP. I. Undue industry influences that distort healthcare research, strategy, expenditure and practice: a review. *European Journal of Clinical Investigation* 2013;43:469-75.
13. Bruix J, Merle P, Granito A, Huang Y-H, Bodoky G, Yokosuka O, *et al.* Efficacy, safety, and health-related quality of life (HRQoL) of regorafenib in patients with hepatocellular carcinoma (HCC) progressing on sorafenib: Results of the international, double-blind phase 3 RESORCE trial. *Annals of Oncology* 2016;S6:LBA28.
14. Bruix J, Tak W-Y, Gasbarrini A, Santoro A, Colombo M, Lim H-Y, *et al.* Regorafenib as second-line therapy for intermediate or advanced hepatocellular carcinoma: Multicentre, open-label, phase II safety study. *European Journal of Cancer* 2013;49:3412-19.
15. Solms A, Reinecke I, Fiala-Buskies S, Keunecke A, Drenth HJ, Bruix J, *et al.* Exposure-response relationship of regorafenib efficacy in patients with hepatocellular carcinoma. *European Journal of Pharmaceutical Sciences* 2017;19 [epub ahead of print].
16. Bruix J, Merle P, Granito A, Huang YH, Bodoky G, Pracht M, *et al.* Analysis of overall survival (OS) by pattern of progression of hepatocellular carcinoma (HCC) during prior sorafenib

treatment in the randomized phase 3 RESORCE trial comparing regorafenib with placebo. *Hepatology International* 2017;11 (1 Supplement 1):S63.

17. Pugh R, Murray-Lyon I, Dawson J, Pietroni M, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *British Journal of Surgery* 1973;60:646-49.

18. Bayer PLC Ltd. RESORCE - Clinical Study Report; 2016.

19. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Seminars in Liver Disease* 2010;30:52-60.

20. Eisenhauer E, Therasse P, Bogaerts J. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European Journal of Cancer* 2009;45:228-47.

21. Llovet JM, Ricci S, Mazzaferro V. Sorafenib in advanced hepatocellular carcinoma. *New England Journal of Medicine* 2008;359:378-90.

22. Cella D, Tulsky D, Gray G, Sarafian B, Linn E, Bonomi A. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *Journal of Clinical Oncology* 1993;11:570-79.

23. Heffernan N, Cella D, Webster K, Odom L, Martone M, Passik S. Measuring health-related quality of life in patients with hepatobiliary cancers: the functional assessment of cancer therapy-hepatobiliary questionnaire. *Journal of Clinical Oncology* 2002;20:2229-39.

24. Pickard A, Neary M, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health and Quality of Life Outcomes* 2007;5:70.

25. Bruix J, Cheng AL, Meinhardt G, Nakajima K, De Sanctis Y, Llovet JM. Prognostic Factors and Predictors of Sorafenib Benefit in Patients With Hepatocellular Carcinoma: Analysis of Two Phase 3 Studies. *Journal of Hepatology* 2017 [epub ahead of print].

26. King J, Palmer J, Johnson P, Ross P, Hubner R, Sumpter K*, et al.* Sorafenib for the Treatment of Advanced Hepatocellular Cancer - a UK Audit. *Clinical Oncology* 2017;29:256-62.

27. Steel J, Eton D, Cella D, Olek M, Carr B. Clinically meaningful changes in health-related quality of life in patients diagnosed with hepatobiliary carcinoma. *Annals of Oncology* 2006;17:304-12.

28. European Medicines Agency. Summary of product characteristics – regorafenib. EMA: London. 2017. In. EMA: London; 2017.

29. Grothey A, Van Cutsem E, Sobrero A, Siena S. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *The Lancet* 2013;381:303-12.

30. Li J, Quin S, Xu R, Yao T, Ma B, Pan H*, et al.* Regorafenib plus best supportive care versus placebo plus best supportive care in Asian patients with previously treated metastatic colorectal cancer (CONCUR): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncology* 2015;16:619-29.

31. Gyawali B, Prasad V. Health policy: Me-too drugs with limited benefits - the tale of regorafenib for HCC. *Nature Reviews Clinical Oncology* 2017;18:18.

32. Frenette CT. The role of regorafenib in hepatocellular carcinoma. *Gastroenterology and Hepatology* 2017;13:122-4.

33. Prasad V, Massey P, Fojo T. Oral anticancer drugs: how limited dosing options and dose reductions may affect outcomes in comparative trials and efficacy in patients. *Journal of Clinical Oncology* 2014;32:1620-29.

34. Bruix J, Raoul J, Sherman M, Mazzaferro V, Bolondi L, Craxi A*, et al.* Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma: subanalyses of a phase III trial. *Journal of Hepatology* 2012;57:821-29.

35. Parikh ND, Singal AG, Hutton DW. Cost effectiveness of regorafenib as second-line therapy for patients with advanced hepatocellular carcinoma. *Cancer* 2017; http://dx.doi.org/10.1002/cncr.30863.

36. National Institute for Health and Clinical Excellence. Regorafenib for previously treated unresectable hepatocellular carcinoma. Final scope. In. NICE: London, UK.; 2017: 1-3.

37. Department of Health. NHS Reference Costs 2015/16. 2016. https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016

38. Curtis L, Burns A. Unit Costs of Health and Social Care 2016.; 2016.

39. Akhtar W, Chung Y. Saving the NHS one blood test at a time. *BMJ Quality Improvement Reports* 2014;2:u204012.w1749.

40. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013; 2013.

41. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the economic evaluation of health care programmes. 4th edition. edn. Oxford, England: Oxford University Press.

42. 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 : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2003;6:9-17.

43. Grimm S BI, Carroll C, Wong R. Cancer Drugs Fund Rapid Review of Nice Guidance Ta189: Sorafenib For The Treatment of Advanced Hepatocellular Carcinoma. 2016. https://www.nice.org.uk/guidance/gid-ta10055/documents/committee-papers (Accessed

9 APPENDICES

Appendix 1: Questions posed to the company and the answers provided regarding the resource use survey undertaken by the company.

During the STA process, but subsequent to the formal clarification process the ERG noticed that there is a key discrepancy in the results of the survey used by the company in its submission. The ERG informed NICE of this and NICE formulated and sent an additional question to the company. (Question 1). The company responded to NICE's question with Answer 1. The ERG were not satisfied with the company's answer and provided a more detailed explanation of the discrepancy (Question 2). The company responded to Question 2 with Answer 2.

Questions 1 and 2, and Answers 1 and 2 are replicated in this Appendix. Following these, the thoughts of the ERG in relation to Answer 2 are provided.

Question 1. "In Table 121 of Appendix 0 in the top right hand cell of the table, clinicians were asked to provide the 'number of admissions per month'. However, the ERG have identified that in all uses of this parameter the number is less than 1. The number of admissions per month must logically be 1 or above"

Answer 1. "The questionnaire asks for the number of admissions per month. Throughout the questionnaire physicians were instructed to enter a decimal if the unit of interest occurred less frequently than once per month. For example, if the frequency was once every three months they were instructed to enter 0.33 (1 divided by 3). On this basis parameter values less than one are logical and in this case indicate that patients on average are hospitalised less than once per month."

Question 2. "The ERG believes that the implementation of the survey resources is unlikely to be consistent with the way in which it was intended. They demonstrate this using the costs of hospitalisation, assuming that the 2015 survey is appropriate. For pre-progression patients having sorafenib the cost of a general ward stay is calculated using four elements:

- P – the proportion of patients requiring hospitalisation ([redacted]
- D – the duration of an average ward stay ([redacted] days)
- H – the number of hospitalisations ([redacted])
- C – cost per bed day (£801)

The model calculates the cost per cycle as P * D * H * C

This formula is only conceptually correct if P is dividing the population into those who are susceptible to hospitalisation and those who are immune. In this instance, H would be therefore only be applied to P.

If, however, P is the proportion of the total population who are hospitalised then H is not needed, unless the model is taking multiple hospitalisations into account.

If multiple hospitalisations are not included the formula should be P*D*C

If multiple hospitalisations are included, then P*D*H*C is correct but H would need to be ≥ 1. (In the submitted model H<1 in all cases)

A similar problem applies for other costs. For example, in cell N84 of the costs sheet it is assumed that for best supportive care, [redacted] of patients do not have INR tests, and that the remaining [redacted] of patients average 0.67 tests.

The instructions to the questionnaire states that the expert should assess the 'average or typical' patient. As such, it is unlikely that they would be answering assuming a proportion of susceptible patients. Can you clarify what the clinicians were intended to be asked? Can you also provide plausible reasons for why P * D * H * C, with H < 1 is correct?"

Answer 2

"The survey question to which this query relates is in the box below. 'A' and 'B' have been added to the table in order to try to make our response clearer.

Acute Care

Q6a Still thinking about a typical 'pre-progression' advanced HCC patients, what proportions of patients receiving treatment or taking no other active treatment (BSC) require each of the following resources as part of acute care?

Table 121. Acute care for 'pre-progression' patients with advanced HCC for patients treated with sorafenib

Acute Care	Average proportions for 'pre-progression' patients	Number of admissions per month
Proportion requiring a hospitalisation (per month)	A	B

The intention of this question was to isolate what has been described above as 'susceptible' patients i.e. the subset of patients who are hospitalised. Looking at this group of patients the questionnaire was structured to allow for multiple hospitalisations as our *a priori* assumption was that for population 'A' the number of admissions 'B' would have a lower bound of 1 (as put forward by the ERG). According to the intention of the questionnaire the appropriate inclusion in the model is P*H*D*C.

The ERG queries whether the questionnaire has been answered as intended given the response to the number of hospitalisations is less than 1.

We believe the response indicates that [redacted] of patients are hospitalised in the pre-progressed health state and that the number of admissions per month relates to this group, as was the intention. We appreciate that for this to be the case the respondents would need to have not seen the time period of 'one month' as indicated in the first column of the table – however, we suspect this is what has happened. We have sought expert clinical opinion regarding this which supports our interpretation (see 'Expert Opinion' overleaf). The responses therefore indicate that [redacted] of patients are 'susceptible' in the pre-progressed health state and that admissions occur less frequently than monthly for this specific population.

We believe it is possible that the survey question immediately preceding question 6a (see box below) may have had an influence. In question 5 respondents were guided to enter decimals when frequencies were less than once per month (i.e. once every three months).

Medical Staff Visits

Q5 Furthermore, when thinking about a typical 'pre-progression' advanced hepatocelluar carcinoma patients, on average how many and which type of physician, nurse and GP visits do they receive per month. If the test is likely to be performed less than once a month enter a decimal e.g. if performed once every 3 months enter 0.333 (1 divided by 3). Please keep in mind that this section is referring to any visits that would be planned (elective).

Table 120. Medical staff visits for 'pre-progression' patients with advanced HCC

Physician visits	Average number of visits (per month) and specialty if required
Pre-progression patients treated with sorafenib	
Specialist visit (e.g. oncologist, gastroenterologist etc.)	
Nurse visit (e.g. clinical nurse specialist, palliative care nurse etc.)	
GP visit	
Other physician visit (please specify)	
Pre-progression patients on BSC	
Specialist visit (e.g. oncologist, gastroenterologist etc.)	
Nurse visit (e.g. clinical nurse specialist, palliative care nurse etc.)	
GP visit	
Other physician visit (please specify)	

Expert opinion

Ideally it would have been possible to contact the original respondents to seek clarification on their answers. However, according to the market research code of practice re-contacting respondents is only allowed if permission is formally provided – such permission was not obtained by the medical research agency. We therefore sought advice from a clinical expert experienced in the management of advanced HCC. Based on their clinical experience they consider that the questionnaire is likely to have been answered as we thought since this is the level of hospitalisation they would expect i.e. [redacted] of patients have [redacted] admissions per month.

Sensitivity analysis

We have presented a sensitivity analysis assuming that [redacted] of patients are hospitalised with an admission frequency of once per month i.e. P*D*C. No changes have been made to the laboratory/radiological tests as the 2015 resource survey asks for the proportion of patients requiring each resource and the frequency 'of these' patients receiving each resource – the implementation in the model is correct."

Table 5. Sensitivity analysis – [redacted] of patients are hospitalised once/month

	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Base case	14,625	0.406	36,050
Sensitivity analysis	15,538	0.406	38,303

The ERG's thoughts on Answer 2.

The ERG does not believe it likely that the post-hoc justification provided by the company regarding how the questions on hospitalisations were filled in by the clinicians is correct. This is for two reasons: (i) it would appear simpler for a clinician to try and estimate the proportion of the whole population that is hospitalised each month, rather than to assume that only a certain proportion of patients could be hospitalised, whilst the remainder would not, and then to estimate a rate of hospitalisation for those susceptible and (ii) clinical advice provided to the ERG states that the risk of hospitalisation would not be zero for any of the population considered. The ERG believe that the sensitivity analysis conducted by the company, where the number of admissions per month (denoted 'B' by the company) are set to 1 are more suitable for decision making than the company's base case analyses.

Appendix 2: Technical appendix – implementation of ERG exploratory analyses

This appendix details the amendments made to the company's model within the ERG's exploratory analyses.

Exploratory analysis 1: Correction of unequivocal model errors and use of alternative unit costs
The values in the following cells were changed:

- "Model Summary" sheet cell O25 changed from no to yes.
- "Model Summary" sheet cell O26 changed from no to yes.
- "Model Summary" sheet cell O27 changed from no to yes.
- "Patient cohort" sheet cell BK30 amended to:
  =IF(IF(effect!$E$329=2,Pat_cohort!BM30,IF(effect!$J$274=1,BF30,IF(effect!$J$274=2,BE30,IF(effect!$J$274=3,BG30,IF(effect!$J$274=4,BH30,IF(effect!$J$274=5,BI30,IF(effect!$J$274=6,BD30,BJ30)))))))>CN30,CN30,IF(effect!$E$329=2,Pat_cohort!BM30,IF(effect!$J$274=1,BF30,IF(effect!$J$274=2,BE30,IF(effect!$J$274=3,BG30,IF(effect!$J$274=4,BH30,IF(effect!$J$274=5,BI30,IF(effect!$J$274=6,BD30,BJ30)))))))
- "Costs" sheet cell L32 changed from 64.6% to 0%.
- "Costs" sheet cell F37 was changed from £138 to £204.11.
- "Costs" sheet cell F43 was changed from £131 to £119.03.
- "Costs" sheet cell F49 was changed from £151 to £162.84.
- "Costs" sheet cell F59 was changed from £238 to £202.70.
- "Live" sheet cell C96 was changed from £1,225.99 to £1,184.11.
- "Live" sheet cell C100 was changed from £1,225.99 to £1,184.11.
- "Live" sheet cell E96 was changed from £1,493.22 to £1,365.07.

In addition, the formulae in the following cells on the "Model_cost" sheet were changed:

- The formulae in cells AC284 to AC523 were amended to refer to "Live" sheet cell E68 rather than Live sheet cell G68
- The formulae in cells AC534 to AC773 were amended to refer to "Live" sheet E69 rather than G69.

All amended cells are highlighted in red in the ERG's revised model.

<u>All subsequent exploratory analyses are based on this amended version of the model.</u>

Exploratory analysis 2: Inclusion of more appropriate general ward bed day cost
Worksheet "Costs" cell F36 was changed from £801 to £572.44.

This amendment can be implemented by entering a value of "yes" into worksheet "Model Summary" cell O37 in the ERG's revised model.

Exploratory analysis 3: Use of full pack dosing
The values in the following cells were changed:

- "Live" sheet cell C36 was changed to 160mg.
- "Live" sheet cell C37 was changed to 160mg.

This amendment can be implemented by entering a value of "yes" into worksheet "Model Summary" cell O38 in the ERG's revised model.

Exploratory analysis 4: Removal of half-cycle correction for drug acquisition costs
Within the worksheet "Model_cost", the following cells were amended:

- The formulae in cells Q34:Q273 which referred to cells Q275:Q514 in the "patient cohort" worksheet were changed to refer to cells Q29:Q270 in the "patient cohort" sheet.
- The formulae in cells R34:R273 which referred to cells U275:U514 in the "patient cohort" sheet were changed to refer to cells U29:U270 in the "patient cohort" sheet.
- The formulae in cells S34:S273 which referred to cells S275:S514 in the "patient cohort" sheet were changed to refer to cells S29:S270 in the "patient cohort" sheet.
- The formula in cells AS34:AS273 which referred to cells AD275:AD514 in the "patient cohort" sheet were changed to refer to cells AD29:AD270 in the "patient cohort" sheet.
- The formula in cells AT34:AT273 which referred to cells AH275:AH514 in the "patient cohort" sheet were changed to refer to cells AH29:AH270 in the "patient cohort" sheet.
- The formula in cells AU34:AU273 which referred to AF275:AF514 in the "patient cohort" sheet were changed to refer to cells AF29:AF270 in the "patient cohort" sheet.

This amendment can be implemented by entering a value of "yes" into worksheet "Model Summary" cell O39 in the ERG's revised model.

Exploratory analysis 5: Use of combined 2007 and 2015 survey costs
Within the "Costs" worksheet, cells H68, L68, O68 and S68 were changed to values of 1.00.

Within the "summary" worksheet, cell O29 was changed from "no" to "yes". Selecting "yes" automatically triggers the amended values described above in the ERG's revised model.

Exploratory analysis 6: Use of independent Weibull functions to model OS

The option selected from the "survival curves" box on the worksheet "effect" was changed from "dependent curves" to "independent curves" for OS. The Weibull model was selected from the relevant drop down box for OS.

Exploratory analysis 7: Use of a fully extrapolated log logistic time to discontinuation curve (patients on treatment at 29th February 2016 censored, with full pack dosing)

Please refer to new worksheet "ERG_TTD." This replaces the company's cost estimates generated in worksheet "KM discontinuation."

- Columns B:E are linked to the log logistic time to treatment discontinuation curve in worksheet "KM discontinuation."
- Columns G:I summarise the probability of being alive, based on the currently selected OS model.
- Columns K:O introduce a logical consistency constraint which ensures that the modelled time to treatment discontinuation curve never exceeds the OS curve.
- Column Q estimates the total drug cost based on the probability of a patient being alive and still on treatment at the beginning of each cycle and the regorafenib acquisition cost per treatment cycle (cell D2).
- Column R discounts the per-cycle treatment cost at a rate of 3.5% per annum.
- The total discounted treatment cost calculated in cell U3 is linked to worksheet "Output" cell I33.

This amendment can be implemented by entering a value of "yes" into worksheet "Model Summary" cells O30 and O42 in the ERG's revised model.

Exploratory analysis 8: ERG's preferred base case (including all individual amendments)

All changes detailed above were implemented together.